SuperStack® 3
Switch Implementation Guide

For units in the SuperStack 3 Switch 4900 Series

http://www.3com.com/

Part No. DUA1770-0BAA05
Rev. 01
Published November 2002

3Com Corporation
5400 Bayfront Plaza
Santa Clara, California
95052-8145

3Com Corporation reserves the right to revise this documentation and to make changes in content from time to time without obligation on the part of 3Com Corporation to provide notification of such revision or change.

3Com Corporation provides this documentation without warranty, term, or condition of any kind, either implied or expressed, including, but not limited to, the implied warranties, terms or conditions of merchantability, satisfactory quality, and fitness for a particular purpose. 3Com may make improvements or changes in the product(s) and/or the program(s) described in this documentation at any time.

If there is any software on removable media described in this documentation, it is furnished under a license agreement included with the product as a separate document, in the hard copy documentation, or on the removable media in a directory file named LICENSE.TXT or !LICENSE.TXT. If you are unable to locate a copy, please contact 3Com and a copy will be provided to you.

UNITED STATES GOVERNMENT LEGEND

If you are a United States government agency, then this documentation and the software described herein are provided to you subject to the following:

All technical data and computer software are commercial in nature and developed solely at private expense. Software is delivered as "Commercial Computer Software" as defined in DFARS 252.227-7014 (June 1995) or as a "commercial item" as defined in FAR 2.101(a) and as such is provided with only such rights as are provided in 3Com's standard commercial license for the Software. Technical data is provided with limited rights only as provided in DFAR 252.227-7015 (Nov 1995) or FAR 52.227-14 (June 1987), whichever is applicable. You agree not to remove or deface any portion of any legend provided on any licensed program or documentation contained in, or delivered to you in conjunction with, this User Guide.

Unless otherwise indicated, 3Com registered trademarks are registered in the United States and may or may not be registered in other countries.

3Com, the 3Com logo, and SuperStack are registered trademarks of 3Com Corporation.

Intel and Pentium are registered trademarks of Intel Corporation. Microsoft, MS-DOS, Windows, and Windows NT are registered trademarks of Microsoft Corporation. Novell and NetWare are registered trademarks of Novell, Inc. UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company, Ltd.

IEEE and 802 are registered trademarks of the Institute of Electrical and Electronics Engineers, Inc.

All other company and product names may be trademarks of the respective companies with which they are associated.

ENVIRONMENTAL STATEMENT

It is the policy of 3Com Corporation to be environmentally-friendly in all operations. To uphold our policy, we are committed to:

Establishing environmental performance standards that comply with national legislation and regulations.

Conserving energy, materials and natural resources in all operations.

Reducing the waste generated by all operations. Ensuring that all waste conforms to recognized environmental standards. Maximizing the recyclable and reusable content of all products.

Ensuring that all products can be recycled, reused and disposed of safely.

Ensuring that all products are labelled according to recognized environmental standards.

Improving our environmental record on a continual basis.

End of Life Statement

3Com processes allow for the recovery, reclamation and safe disposal of all end-of-life electronic components.

Regulated Materials Statement

3Com products do not contain any hazardous or ozone-depleting material.

Environmental Statement about the Documentation

The documentation for this product is printed on paper that comes from sustainable, managed forests; it is fully biodegradable and recyclable, and is completely chlorine-free. The varnish is environmentally-friendly, and the inks are vegetable-based with a low heavy-metal content.

CONTENTS

About This Guide

This guide describes the features of the SuperStack® 3 Switch 4900 Series and outlines how to use these features to optimize the performance of your network.

Most features detailed in this guide are common to all Switches in the 4900 Series. Refer to the Management Quick Reference Guide that accompanies your Switch for details of the specific features your Switch supports.

This guide is intended for the system or network administrator who is responsible for configuring, using, and managing the Switches. It assumes a working knowledge of local area network (LAN) operations and familiarity with communication protocols that are used to interconnect LANs.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM that accompanies your Switch or on the 3Com Web site.

If release notes are shipped with your product and the information there differs from the information in this guide, follow the instructions in the release notes.

Most user guides and release notes are available in Adobe Acrobat Reader Portable Document Format (PDF) or HTML on the 3Com World Wide Web site:

http://www.3com.com/

Conventions

Table 1 and Table 2 list conventions that are used throughout this guide.

Table 1 Notice Icons

Icon	Notice Type	Description
	Information note	Information that describes important features or instructions
	Caution	Information that alerts you to potential loss of data or potential damage to an application, system, or device
	Warning	Information that alerts you to potential personal injury

Table 2 Text Conventions

Convention	Description
`Screen displays`	This typeface represents information as it appears on the screen.
`Syntax`	The word "syntax" means that you must evaluate the syntax provided and then supply the appropriate values for the placeholders that appear in angle brackets. Example: To change your password, use the following syntax: `system password <password>` In this example, you must supply a password for <password>.
`Commands`	The word "command" means that you must enter the command exactly as shown and then press Return or Enter. Commands appear in bold. Example: To display port information, enter the following command: **`bridge port detail`**
The words "enter" and "type"	When you see the word "enter" in this guide, you must type something, and then press Return or Enter. Do not press Return or Enter when an instruction simply says "type."
Keyboard key names	If you must press two or more keys simultaneously, the key names are linked with a plus sign (+). Example: Press Ctrl+Alt+Del
Words in *italics*	Italics are used to: ■ Emphasize a point. ■ Denote a new term at the place where it is defined in the text. ■ Identify menu names, menu commands, and software button names. Examples: From the *Help* menu, select *Contents*. Click *OK*.

Related Documentation

In addition to this guide, each Switch documentation set includes the following:

- *Getting Started Guide*

  This guide contains:

  - all the information you need to install and set up the Switch in its default state
  - information on how to access the management software to begin managing the Switch.

- *Management Interface Reference Guide*

  This guide contains information about the web interface operations and CLI (command line interface) commands that enable you to manage the Switch. It contains an explanation for each command and the different parameters available. It is supplied in HTML format on the CD-ROM supplied with your Switch, or the 3Com Web site.

- *Management Quick Reference Guide*

  This guide contains:

  - a list of the features supported by the Switch
  - a summary of the web interface operations and CLI commands that enable you to manage the Switch.

- *Release Notes*

  These notes provide information about the current software release, including new features, modifications, and known problems.

In addition, there are other publications you may find useful:

- Documentation accompanying the Expansion Modules.
- Documentation accompanying the Advanced Redundant Power System.

Documentation Comments

Your suggestions are very important to us. They will help make our documentation more useful to you. Please e-mail comments about this document to 3Com at:

pddtechpubs_comments@3com.com

Please include the following information when contacting us:

- Document title
- Document part number (on the title page)
- Page number (if appropriate)

Example:

- SuperStack 3 Switch Implementation Guide
- Part number: DUA1770-0BAA0x
- Page 25

Please note that we can only respond to comments and questions about 3Com product documentation at this e-mail address. Questions related to technical support or sales should be directed in the first instance to your network supplier.

Product Registration

You can now register your SuperStack 3 Switch on the 3Com web site:

www.3com.com/register

1 SWITCH FEATURES OVERVIEW

This chapter contains introductory information about the SuperStack® 3 Switch 4900 Series management software and supported features. It covers the following topics:

What is Management Software?

Your Switch can operate in its default state. However, to make full use of the features offered by the Switch, and to change and monitor the way it works, you have to access the management software that resides on the Switch. This is known as managing the Switch.

Managing the Switch can help you to improve its efficiency and therefore the overall performance of your network.

There are several different methods of accessing the management software to manage the Switch. These methods are explained in Chapter 3 of the Getting Started Guide that accompanies your Switch.

Switch Features Explained

The management software provides you with the capability to change the default state of some of the Switch features. This section provides a brief overview of these features — their applications are explained in more detail later in this guide.

For a list of the features supported by your Switch, please refer to the Management Quick Reference Guide that accompanies your Switch.

Aggregated Links

Aggregated links are connections that allow devices to communicate using up to four links in parallel. Aggregated links provide two benefits:

- They can potentially double, triple or quadruple the bandwidth of a connection.
- They can provide redundancy — if one link is broken, the other links share the traffic for that link.

Your Switch supports the IEEE 802.3ad-2000 standard Link Aggregation Control Protocol (LACP) which means that, if LACP is enabled, your Switch can automatically configure its aggregated link without any manual intervention.

For more information about aggregated links and LACP, see Chapter 2 "Optimizing Bandwidth".

Auto-negotiation

Auto-negotiation allows ports to auto-negotiate port speed, duplex-mode (only at 10 Mbps and 100 Mbps) and flow control. When auto-negotiation is enabled (default), a port "advertises" its maximum capabilities — these capabilities are by default the parameters that provide the highest performance supported by the port.

1000BASE-SX and GBIC ports do not support auto-negotiation of port speed.

Ports operating at 1000 Mbps only support full duplex mode.

For details of the auto-negotation features supported by your Switch, please refer to the Management Quick Reference Guide that accompanies your Switch.

Duplex

Full duplex mode allows packets to be transmitted and received simultaneously and, in effect, doubles the potential throughput of a link.

Flow Control

All Switch ports support flow control, which is a mechanism that minimizes packet loss during periods of congestion on the network.

Flow control is supported on ports operating in half duplex mode, and is implemented using the IEEE Std 802.3-2002 (incorporating 802.3x) on ports operating in full duplex mode.

Smart Auto-sensing

Smart auto-sensing allows auto-negotiating multi-speed ports, such as 10/100 Mbps or 10/100/1000 Mbps, to monitor and detect high error rates, or problems in the "physical" interconnection to another port. The port reacts accordingly by tuning the link from its higher speed to the lower supported speed to provide an error-free connection to the network.

1000BASE-SX and GBIC ports do not support smart auto-sensing.

For more information about auto-negotiation and port capabilities, see Chapter 2 "Optimizing Bandwidth".

Multicast Filtering

Multicast filtering allows the Switch to forward multicast traffic to only the endstations that are part of a predefined multicast group, rather than broadcasting the traffic to the whole network.

The multicast filtering system supported by your Switch uses IGMP (Internet Group Management Protocol) snooping to detect the endstations in each multicast group to which multicast traffic should be forwarded.

For more information about multicast filtering, see Chapter 3 "Using Multicast Filtering".

Resilient Links

The resilient link feature enables you to protect critical links and prevent network downtime should those links fail. Setting up resilient links ensures that if a main communication link fails, a standby duplicate link automatically takes over the task of the main link. Each main and standby link pair is referred to as a resilient link pair.

Resilient links are a simple method of creating redundancy that provides you with a fast reaction to link failure. Resilient links are quick to set up, you have full control over their configuration, and the port at the other end of the resilient link does not have to support any resilience feature.

For more information about resilient links, see Chapter 4 "Using Resilience Features".

Spanning Tree Protocol and Rapid Spanning Tree Protocol

Spanning Tree Protocol (STP) and Rapid Spanning Tree Protocol (RSTP) are bridge-based protocols that make your network more resilient to link failure and also provide protection from network loops — one of the major causes of broadcast storms.

STP/RSTP allows you to implement alternative paths for network traffic in the event of path failure and uses a loop-detection process to:

- Discover the efficiency of each path.
- Enable the most efficient path.
- Disable the less efficient paths.
- Enable one of the less efficient paths if the most efficient path fails.

RSTP is an enhanced version of the STP feature and is enabled by default. RSTP can restore a network connection quicker than the legacy STP feature. RSTP can detect if it is connected to a legacy device that only supports IEEE 802.ID STP and will automatically downgrade to STP on that particular port.

STP conforms to the IEEE Std 802.1D, 1998 Edition and RSTP conforms to the IEEE Std 802.1w-2001.

For more information about STP, see Chapter 4 "Using Resilience Features".

Switch Database

The Switch Database is an integral part of the Switch and is used by the Switch to determine if a packet should be forwarded, and which port should transmit the packet if it is to be forwarded.

For more information about the Switch Database, see Chapter 5 "Using the Switch Database".

Traffic Prioritization

Traffic prioritization allows time-sensitive and system-critical data, such as digital video and network-control signals, to be transferred smoothly and with minimal delay over a network. This data is assigned a high priority by the transmitting endstation and traffic prioritization allows high priority data to be forwarded through the Switch without being obstructed by lower priority data.

The prioritization works by using the multiple traffic queues that are present in the hardware of the Switch — high priority data is forwarded

on a different queue from lower priority data, and is given preference over the lower priority data.

This system is compatible with the relevant sections of the IEEE 802.ID, 1998 Edition.

For more information about 802.1D and traffic prioritization, see Chapter 6 "Using Traffic Prioritization".

Quality of Service

Traffic prioritization can be taken one step further by using the Quality of Service (QoS) feature. Policy-based Quality of Service (QoS) enables you to specify service levels for different traffic classifications. This enables you to prioritize particular applications or traffic types.

The Switch uses a policy-based QoS mechanism. By default, all traffic is assigned the "normal" QoS policy profile. If needed, you can create other QoS policy profiles and apply them to different traffic types so that they have different priorities across the network.

For more information about Quality of Service, see Chapter 6 "Using Traffic Prioritization".

RMON Remote Monitoring (RMON) is a system that allows you to monitor LANs remotely. The Switch software continually collects statistics about the LAN segments connected to the Switch. If you have a management workstation with an RMON management application, the Switch can transfer these statistics to your workstation on request or when a pre-defined threshold is exceeded.

Event Notification

You can configure your Switch to send you notification when certain events occur. You can receive notification via email, SMS (Short Message Server), or pager.

For more information about RMON, see Chapter 7 "Status Monitoring and Statistics".

Trace Route

Trace Route allows you to trace the route of IP packets through your network from a local device to a remote destination. This is useful to assist in monitoring or troubleshooting your network.

For more information about Trace Route, see Chapter 7 "Status Monitoring and Statistics".

VLANs

A Virtual LAN (VLAN) is a flexible group of devices that can be located anywhere in a network, but which communicate as if they are on the same physical segment. With VLANs, you can segment your network without being restricted by physical connections — a limitation of traditional network design. As an example, with VLANs you can segment your network according to:

- Departmental groups
- Hierarchical groups
- Usage groups

For more information about VLANs, see Chapter 8 "Setting Up Virtual LANs".

Automatic IP Configuration

By default the Switch tries to configure itself with IP information without requesting user intervention. The Switch uses the following industry standard methods in turn to allocate the Switch IP information:

- Dynamic Host Configuration Protocol (DHCP)
- Auto-IP — the Switch will configure itself with its default IP address 169.254.100.100 if it is operating in a standalone mode, and/or no other Switches on the network have this IP address. If this default IP address is already in use on the network then the Switch detects this and configures itself with an IP address in the range 169.254.1.0 to 169.254.254.255.
- Bootstrap Protocol (BOOTP)

For more information about how the automatic IP configuration feature works, see Chapter 9 "Using Automatic IP Configuration".

IP Routing

IP Routing is a method for distributing traffic throughout an IP network. It is used to join LANs at the network layer, that is Layer 3 of the OSI (Open Systems Interconnection) model. Routers are used to:

- Connect enterprise networks.
- Connect subnetworks (or client/server networks) to the main enterprise network.

For more information about Layer 3 Routing, see Chapter 10 "IP Routing".

Webcache Support

Webcache support allows your Switch to detect and redirect HTTP web traffic to a local Webcache. Users can then access frequently used Web pages stored locally on the Webcache — this allows your network to operate more efficiently and reduces WAN network traffic.

For more information about Webcache Support, see Chapter 11 "Using Webcache Support".

Broadcast Storm Control

Broadcast Storm Control is a system that monitors the level of broadcast traffic on that port. If the broadcast traffic level rises to a pre-defined number of frames per second (threshold), the broadcast traffic on the port is blocked until the broadcast traffic level drops below the threshold. This system prevents the overwhelming broadcast traffic that can result from network equipment which is faulty or configured incorrectly. Broadcast Storm Control is enabled by default.

Switch Management Login

In addition to the standard method of Switch management login which uses the local Switch database to authenticate login attempts, if you have a RADIUS server on your network you can use this to centrally authenticate all login attempts to all Switches on your network that support the RADIUS protocol. Using a RADIUS server greatly reduces the need for maintenance on network devices as all the login maintenance can be done and controlled centrally and also reduces the risk of security lapses.

For more information about Switch Management Login and RADIUS authentication, see Chapter 12 "Making Your Network Secure".

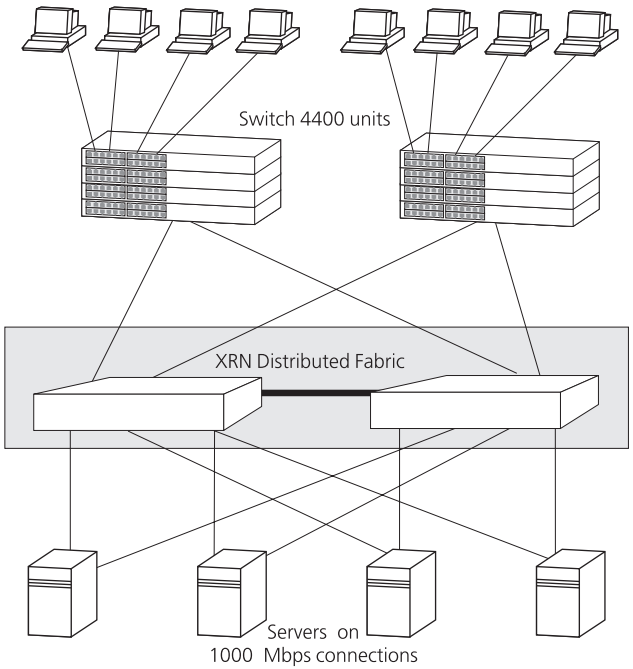
3Com XRN Technology

XRN (eXpandable Resilient Network) is a 3Com technology built into the software and hardware of your Switch that allows you to interconnect Switches to create a Distributed Fabric. The interconnection enables two Switches to behave as a single unit for management and Layer 2 and Layer 3 switching purposes.

This technology provides a highly resilient core around which you can build your network.

Figure 1 shows a simple network configuration using XRN Technology.

Figure 1 XRN Technology Network Example



For more information about XRN Technology and how to implement it in your network, see Chapter 13 "3Com XRN Technology".

2 OPTIMIZING BANDWIDTH

There are many ways you can optimize the bandwidth on your network and improve network performance. If you utilize certain Switch features you can provide the following benefits to your network and end users:

- Increased bandwidth
- Quicker connections
- Faster transfer of data
- Minimized data errors
- Reduced network downtime

Port Features

The default state for all the features detailed below provides the best configuration for most users. *In normal operation, you do not need to alter the Switch from its default state.* However, under certain conditions you may wish to alter the default state of these ports, for example, if you are connecting to old equipment that does not comply with the IEEE 802.3x standard.

Duplex

Full duplex allows packets to be transmitted and received simultaneously and, in effect, doubles the potential throughput of a link. Half duplex only allows packets to be transmitted or received at any one time.

To communicate effectively, both devices at either end of a link *must* use the same duplex mode. If the devices at either end of a link support auto-negotiation, this is done automatically. If the devices at either end of a link do not support auto-negotiation, both ends must be manually set to full duplex or half duplex accordingly.

Ports operating at 1000 Mbps support full duplex mode only.

Flow Control

All Switch ports support flow control, which is a mechanism that prevents packet loss during periods of congestion on the network. Packet loss is caused by one or more devices sending traffic to an already overloaded port on the Switch. Flow control prevents packet loss by inhibiting the transmitting port from generating more packets until the period of congestion ends.

Flow control is implemented using the IEEE Std 802.3-2002 (incorporating 802.3x) for ports operating in full duplex mode, and Intelligent Flow Management (IFM) for ports operating in half duplex mode.

Auto-negotiation

Auto-negotiation allows ports to automatically determine the best port speed, duplex-mode (only at 10 Mbps and 100 Mbps) and flow control. When auto-negotiation is enabled (default), a port "advertises" its maximum capabilities — these capabilities are by default the parameters that provide the highest performance supported by the port.

You can modify the capabilities that a port "advertises" on a per port basis, dependant on the type of port.

1000BASE-SX and GBIC ports do not support auto-negotiation of port speed.

Ports operating at 1000 Mbps support full duplex mode only.

If auto-negotiation is disabled, the auto-MDIX feature does not operate on the ports. Therefore the correct cables, that is, cross-over or straight-through need to be used. For more information, see the Getting Started Guide that accompanies your Switch.

Conditions that affect auto-negotiation:

- Ports at both ends of the link must be set to auto-negotiate.
- 1000BASE-SX ports support auto-negotiation, however, the standard defines that 1000BASE-SX can only operate at 1000 Mbps, full duplex mode, so they can only auto-negotiate flow control.

Smart Auto-sensing

Smart auto-sensing allows auto-negotiating multi-speed ports, such as 100/1000 Mbps, to monitor and detect a high error rate on a link, or a problem in the "physical" interconnection to another port and react

accordingly. In other words, auto-negotiation may "agree" upon a configuration that the link cannot sustain; smart auto-sensing can detect this and adjust the link accordingly.

For example, smart auto-sensing can detect network problems, such as an unacceptably high error rate or a poor quality cable. If both ends of the link support 100/1000 Mbps auto-negotiation, then auto-sensing tunes the link to 100 Mbps to provide an error-free 100 Mbps connection to the network.

An SNMP Trap is sent every time a port is down-rated to a lower speed.

Conditions that affect smart auto-sensing:

- Smart auto-sensing will not operate on links that do not support auto-negotiation, or on links where one end is at a fixed speed. The link will reset to the higher speed of operation when the link is lost or the unit is power cycled.
- Smart auto-sensing can only be configured for the whole Switch and not on a per port basis.

1000BASE-SX and GBIC ports do not support smart auto-sensing.

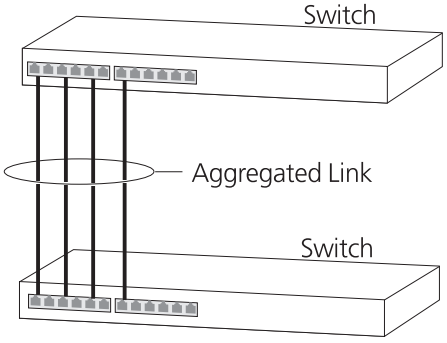
Aggregated Links

Aggregated links are connections that allow devices to communicate using up to four links per Switch in parallel. Aggregated links provide the following benefits:

- They can potentially multiply the bandwidth of a connection. The capacity of the multiple links is combined into one logical link.
- They can provide redundancy — if one link is broken, the other links share the traffic for that link.

Figure 2 shows two Switches connected using an aggregated link containing four member links. If all ports on both Switch units are configured as 100BASE-TX and they are operating in full duplex, the potential maximum bandwidth of the connection is 800 Mbps.

Figure 2 Switch units connected using an aggregated link



3Com recommends that you use IEEE 802.3ad LACP automatic aggregations rather than manual aggregations to ensure maximum resilience on your network. Using manual aggregations to connect to an XRN Distributed Fabric could result in traffic loss and network problems if the XRN Interconnect fails. By default, LACP is disabled on all Switch ports.

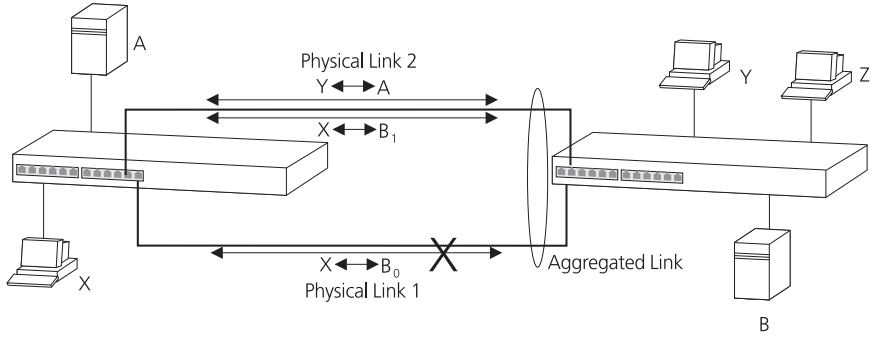
How 802.3ad Link Aggregation Operates

Your Switch supports IEEE Std 802.3-2002 (incorporating 802.3ad) aggregated links which use the Link Aggregation Control Protocol (LACP). LACP provides automatic, point-to-point redundancy between two devices (switch-to-switch or switch-to-server) that have full duplex connections operating at the same speed.

If LACP is enabled on all Switch ports, your Switch will detect if there is more than one connection to another device and will automatically create an aggregated link consisting of those links.

If a member link in an aggregated link fails, the traffic using that link is dynamically reassigned to the remaining member links in the aggregated link. Figure 3 shows the simplest case: two member links, that is the physical links, form an aggregated link. In this example, if link 1 fails, the data flow between X and B is remapped to physical link 2. The re-mapping occurs as soon as the Switch detects that a member link has failed — almost instantaneously. As a result, aggregated link configurations are extremely resilient and fault-tolerant.

Figure 3 Dynamic Reassignment of Traffic Flows



The key benefits of 802.3ad link aggregation are:

- Automatic configuration — network management does not need to be used to manually aggregate links.
- Rapid configuration and reconfiguration — approximately one to three seconds.
- Compatibility — non-802.3ad devices can interoperate with 802.3ad enabled devices. However, you will need to manually configure the aggregated links as LACP will not be able to automatically detect and form an aggregation with a non-802.3ad device.
- The operation of 802.3ad can be configured and managed via network management.

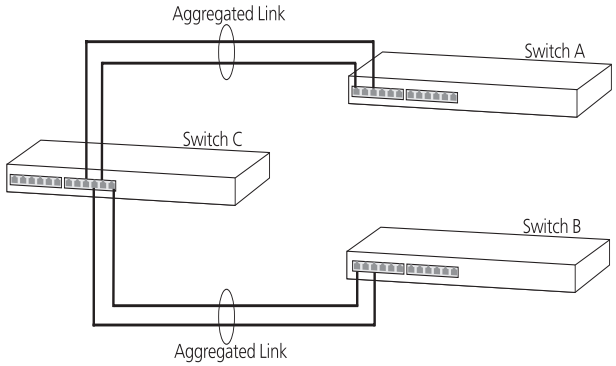
Implementing 802.3ad Aggregated Links

LACP can be enabled or disabled on a per port basis. You can implement 802.3ad aggregated links in three ways:

- Manual Aggregations — You can manually add and remove ports to and from an aggregated link via Web or CLI commands. However, if a port has LACP enabled, and if a more appropriate or correct automatic membership is detected by LACP, it will override the manual configuration.

  For example, in Figure 4, if a port on Switch C is physically connected to Switch B, but you manually configure the port on Switch C to be a member of an aggregated link for Switch A in error, LACP (if it is enabled) will detect this and place the port in the aggregated link for Switch B, thus overriding the manual configuration.

Figure 4 Aggregated Link — Example



- LACP Pre-Configured Aggregations — If you need to know which aggregated link is associated with which device in your network you can use a LACP pre-configured aggregation. This allows you to manually configure the MAC address of a particular partner device (called the partner ID) against a specified aggregated link. LACP will then automatically determine the port membership for that aggregated link.

  The aggregated link may be manually configured with appropriate configuration settings, such as VLAN membership, to match the partner device.

- LACP Automatic Aggregations — If LACP detects at least two active ports sharing the same partner device, and if no matching pre-configured aggregated links exist, LACP will automatically assign a free un-configured aggregated link to form an aggregated link with the partner device. The aggregated link will inherit its configuration from the first port originally detected against the partner device.

  If you have an existing single port connection between two devices, this automatic behavior allows quick and easy addition of extra bandwidth by simply adding an extra physical link between the units.

The Spanning Tree costs for a port running LACP is the cost assigned for an aggregated link running at that speed. As required by the IEEE Std 802.3-2002 (incorporating 802.3ad), no changes in cost are made according to the number of member links in the aggregated link.

If an automatic aggregated link (created by LACP) contains ports with different VLAN membership, the aggregated link will inherit the VLAN memberhip of the first port that comes up in the aggregated link. It will override any pre-defined VLAN membership for the aggregated link. You

therefore need to ensure that prior to the aggregated link forming, every individual port that will be in the aggregated link has the required VLAN memberships configured.

Aggregated Links and Your Switch

- When any port is assigned to an existing aggregated link (either manually or via LACP) it will adopt the configuration settings of the aggregated link. When a port leaves an aggregated link its original configuration settings are restored.
- A maximum of thirteen active aggregated links can be created. A maximum of four ports may be added manually to any individual aggregation, but any number may join automatically via LACP. There are however a few points to consider:
  - Switch 4900 Series and Switches 4050/4060 — The Switch only supports transmission of traffic on a maximum of four ports in any individual aggregation. Any extra ports will be used for reception of traffic only.
  - If multiple links are connected between a unit and more than thirteen other devices, only thirteen of the devices will be assigned to aggregated links. The remaining devices will each only have one link made *active*, that is, passing data. All other links will be made *inactive* to prevent loops occuring.

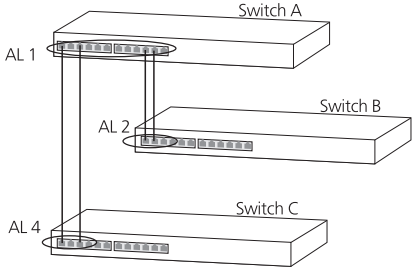
    LACP detects if one of the existing thirteen aggregated links is removed and will then automatically assign one of the remaining devices to the aggregated link that has become free.
- When multiple links of different speed connect two devices only the highest speed links will be aggregated. The other links will be held in a standby state until there is a problem with a higher speed link(s). The lower speed link(s) will then become active.
- Note that resilient links must be disabled on any port that is to become part of an aggregated link. It is not possible to configure resilient links on a port that is a member of an aggregated link, and vice versa.
- A LinkUp / LinkDown trap will only be sent for individual links. The Traps will not be sent for an aggregation.

When setting up an aggregated link, note that:

- The ports at both ends of a member link must be configured as members of an aggregated link, if you are manually configuring aggregated links.

- A member link port can only belong to one aggregated link.
- The member link ports can be mixed media, that is fiber and/or twisted pair ports within the same aggregated link.
- The member link ports can have different port configurations within the same aggregated link, that is, auto-negotiation, port speed, and duplex mode. However, note the following:
  - To be an active participant in an aggregated link the member link ports must operate in full duplex mode. (If a member link port does not operate in full duplex mode it can still be a member of an aggregated link but it will never be activated.)
  - If ports of a different speed are aggregated together, the higher speed links carry the traffic. The lower speed links only carry the traffic if the higher speed links fail.
- Aggregated links and resilient links are mutually exclusive, that is, you cannot have both these features operating on the same ports.
- Member links must retain the same groupings at both ends of an aggregated link. For example, the configuration in Figure 5 will not work as Switch A has one aggregated link defined whose member links are then split between two aggregated links defined on Switches B and C. Note that this illegal configuration could not occur if LACP is enabled.

Figure 5 An illegal aggregated link configuration



To make this configuration work you need to have two aggregated links defined on Switch A, one containing the member links for Switch B and the other containing those for Switch C. Alternatively, if Switches B and C were part of an XRN Distributed Fabric, this configuration would work.

When using an aggregated link, note that:

- To gather statistics about an aggregated link, you must add together the statistics for each port in the aggregated link.

- If you wish to disable a single member link of an aggregated link, you must first physically remove the connection to ensure that you do not lose any traffic, before you disable both ends of the member link separately. If you do this, the traffic destined for that link is distributed to the other links in the aggregated link.

  If you do not remove the connection and only disable one end of the member link port, traffic is still forwarded to that port by the aggregated link port at the other end. This means that a significant amount of traffic may be lost.

- Before removing an entire aggregated link, you must disable all the aggregated link ports or disconnect all the links, except one — if you do not, a loop may be created.

- When manually creating an aggregated link between two devices, the ports in the aggregated link must not be physically connected together until the aggregated link has been correctly configured at both ends of the link. Failure to configure the aggregated link at both ends before physically connecting the ports can result in a number of serious network issues such as lost packets and network loops.

Traffic Distribution and Link Failure on Aggregated Links

To maximize throughput, all traffic is distributed across the individual links that make up an aggregated link. Therefore, when a packet is made available for transmission down an aggregated link, a hardware-based traffic distribution mechanism determines which particular port in the link should be used. The mechanism may use the MAC address, IP address, or a combination of both dependent upon the mode of operation. The traffic is distributed among the member links as efficiently as possible.

To avoid the potential problem of out-of-sequence packets (or "packet re-ordering"), the Switch ensures that all the conversations between a given pair of endstations will pass through the same port in the aggregated link. Single-to-multiple endstation conversations, on the other hand, may still take place over different ports.

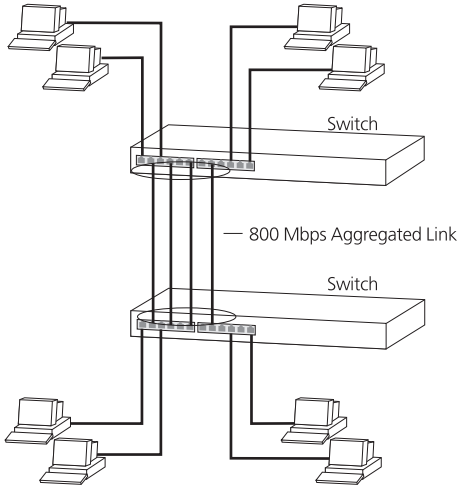
If the link state on any of the ports in an aggregated link becomes inactive due to link failure, then the Switch will automatically redirect the aggregated link traffic to the remaining ports. Aggregated links therefore provide built-in resilience for your network.

The Switch also has a mechanism to prevent the possible occurrence of packet re-ordering when a link recovers too soon after a failure.

Aggregated Link — Manual Configuration Example

The example shown in Figure 6 illustrates an 800 Mbps aggregated link between two Switch units, (that is, each port is operating at 100 Mbps, full duplex).

Figure 6 An 800 Mbps aggregated link between two Switch units



To manually set up this configuration:

1. Prepare ports 2, 4, 6 and 8 on the upper Switch for aggregated links. To do this:
   - **a** Check that the ports have an identical configuration using your preferred management interface.
   - **b** Add the ports 2, 4, 6 and 8 on the specified unit to the aggregated link.
2. Prepare ports 2, 4, 6 and 8 on the lower Switch for aggregated links. To do this:
   - **a** Check that the ports have an identical configuration using your preferred management interface.
   - **b** Add the ports 2, 4, 6 and 8 on the specified unit to the aggregated link.
3. Connect port 2 on the upper Switch to port 2 on the lower Switch.
4. Connect port 4 on the upper Switch to port 4 on the lower Switch.
5. Connect port 6 on the upper Switch to port 6 on the lower Switch.
6. Connect port 8 on the upper Switch to port 8 on the lower Switch.

3 USING MULTICAST FILTERING

Multicast filtering improves the performance of networks that carry multicast traffic.

This chapter explains multicasts, multicast filtering, and how multicast filtering can be implemented on your Switch. It covers the following topics:

- What is an IP Multicast?
- Multicast Filtering
- IGMP Multicast Filtering

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM supplied with your Switch or on the 3Com Web site.

What is an IP Multicast?

A *multicast* is a packet that is intended for "one-to-many" and "many-to-many" communication. Users explicitly request to participate in the communication by joining an endstation to a specific multicast group. If the network is set up correctly, a multicast will only be sent to an endstation or a subset of endstations in a LAN, or VLAN, that belong to the relevant multicast group.

Multicast group members can be distributed across multiple subnetworks; thus, multicast transmissions can occur within a campus LAN or over a WAN. In addition, networks that support IP multicast send only *one* copy of the desired information across the network until the delivery path that reaches group members diverges. It is only at these points that multicast packets are replicated and forwarded, which makes efficient use of network bandwidth.

A multicast packet is identified by the presence of a multicast group address in the destination address field of the packet's IP header.

Benefits of Multicast

The benefits of using IP multicast are that it:

- Enables the simultaneous delivery of information to many receivers in the most efficient, logical way.
- Reduces the load on the source (for example, a server) because it does not have to produce multiple copies of the same data.
- Makes efficient use of network bandwidth and scales well as the number of participants or collaborators expands.
- Works with other IP protocols and services, such as Quality of Service (QoS).

There are situations where a multicast approach is more logical and efficient than a unicast approach. Application examples include distance learning, transmitting stock quotes to brokers, and collaborative computing.

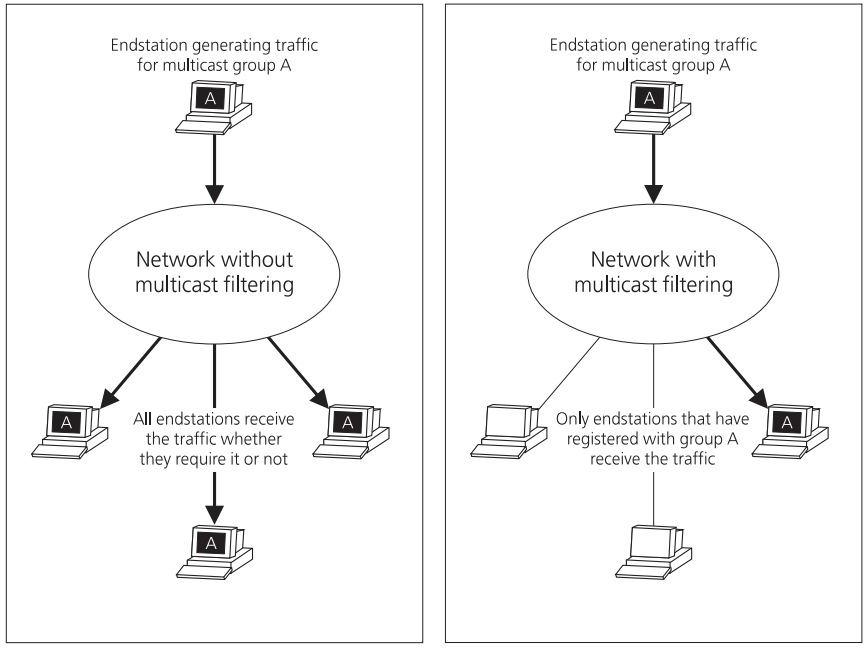
A typical use of multicasts is in video-conferencing, where high volumes of traffic need to be sent to several endstations simultaneously, but where broadcasting that traffic to all endstations would seriously reduce network performance.

Multicast Filtering

Multicast filtering is the system by which endstations only receive multicast traffic if they register to join specific multicast groups. With multicast filtering, network devices only forward multicast traffic to the ports that are connected to registered endstations.

Figure 7 shows how a network behaves without multicast filtering and with multicast filtering.

Figure 7 The effect of multicast filtering



Multicast Filtering and Your Switch

Your Switch provides automatic multicast filtering support using IGMP (Internet Group Management Protocol) Snooping.

Snooping Mode

Snooping Mode allows your Switch to forward multicast packets only to the appropriate ports. The Switch "snoops" on exchanges between endstations and an IGMP device, typically a router, to find out the ports that wish to join a multicast group and then sets its filters accordingly.

The Switch 4900 Series is compatible with any device that conforms to the IGMP v2 protocol. The Switch 4900 Series does not support IGMP v3. If you have an IGMP v3 network, you should disable IGMP snooping for all Switch units in the Distributed Fabric using the **`snoopMode`** *command on the command line interface IGMP menu.*

IGMP Multicast Filtering

IGMP is the protcol that all IP-supporting network devices use to register endstations with multicast groups. It can be used on all LANs and VLANs that contain a multicast capable IP router and on other network devices which support IP.

IGMP multicast filtering works as follows:

1. The IP router (or querier) periodically sends *query* packets to all the endstations in the LANs or VLANs that are connected to it.

   If your network has more than one IP router, then the one with the lowest IP address becomes the querier. The Switch can be the IGMP querier and will become so if its own IP address is lower than that of any other IGMP queriers connected to the LAN or VLAN.

2. When an IP endstation receives a query packet, it sends a *report* packet back that identifies the multicast group that the endstation would like to join.

3. When the report packet arrives at a port on a Switch with *IGMP multicast learning* enabled, the Switch learns that the port is to forward traffic for the multicast group and then forwards the packet to the router.

4. When the router receives the report packet, it registers that the LAN or VLAN requires traffic for the multicast groups.

5. When the router forwards traffic for the multicast group to the LAN or VLAN, the Switch units only forward the traffic to ports that received a report packet.

Enabling IGMP Multicast Learning

You can enable or disable multicast learning and IGMP querying using the `snoopMode` command on the CLI or the web interface. For more information about enabling IGMP multicast learning, please refer to the Management Interface Reference Guide supplied on the 3Com Web site.

If IGMP multicast learning is not enabled then IP multicast traffic is always forwarded, that is, it floods the network.

For information about configuring IGMP functionality on an endstation, refer to the user documentation supplied with your endstation or the endstation's Network Interface Card (NIC).

4 USING RESILIENCE FEATURES

Setting up resilience on your network helps protect critical links against failure, protects against network loops, and reduces network downtime to a minimum.

This chapter explains the software features supported by the Switch that provide resilience for your network. It covers the following topics:

- Resilient Links
- Spanning Tree Protocol (STP)
- Rapid Spanning Tree Protocol (RSTP) — an enhanced version of the STP feature.

Using 3Com XRN Technology to Interconnect two Switches to create a Distributed Fabric provides the maximum level of resilience for your network. For more information on implementing XRN Technology on your network and how it interacts with the features described in this chapter, refer to Chapter 13 "3Com XRN Technology".

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM supplied with your Switch or on the 3Com Web site.

Resilience Feature Overview

Resilient links and STP/RSTP cannot both be used on the network at the same time. Table 3 lists the key differences between each feature, so you can evaluate the benefits of each to determine which feature is most suitable for your network.

Table 3 Resilient Links and Spanning Tree Protocols — Key Differences

Resilient Links	Spanning Tree Protocol	Rapid Spanning Tree Protocol
User configures each Switch separately.	User enables STP on each Switch.	RSTP is enabled by default.
Manual configuration.	Automatic configuration.	Automatic configuration.
Within 5 seconds restores an active connection from a standby link.	Up to 30 second delay on link failure to restoring a network connection.	Within 5 seconds restores a network connection.

3Com recommends that you use the Rapid Spanning Tree Protocol feature (default enabled) to provide optimum performance for your network and ease of use.

RSTP provides the same functionality as STP. For details on how the two systems differ, see "How RSTP Differs to STP" *on* page 45.

The Switch also supports aggregated links which increase bandwidth and also provide resilience against individual link failure. Aggregated links will operate with STP/RSTP enabled, but will not operate on ports that are part of a resilient link pair. For more information, see Aggregated Links on page 25.

What are Resilient Links?

The resilient link feature enables you to protect critical links and prevent network downtime if those links fail. A resilient link is comprised of a *resilient link pair* containing a main link and a standby link. If the main link fails, the standby link quickly and automatically takes over the task of the main link and becomes the "active link".

The resilient link pair is defined by specifying a main port and a standby port at one end of the link. During normal operation, the main port is enabled and the standby port is disabled. If the main link fails, the main port is disabled and the standby port is enabled. If the main link becomes operational, you can then re-enable the main port and disable the standby port again.

There are two user configurable modes of operation for resilient links:

- Symmetric (default) — the standby link remains as the active link even if the main link resumes normal operation.
- Switchback — the standby link continues as the active link until the main link resumes normal operation. The active link then switches back from the standby link to the main link.

When setting up resilient links, note the following:

- Resilient link pairs cannot be set up if the Switch has the Spanning Tree Protocol (STP) or Rapid Spanning Tree Protocol (RSTP) enabled.
- A resilient link pair must only be defined at one end of the link.
- A resilient link pair can only be set up if:
  - The ports use the same VLAN tagging system (IEEE 802.1Q-1998 standard tagging).
  - Neither of the ports are part of an aggregated link.
  - Neither of the ports belong to another resilient link pair.
- The port state of ports in a resilient link pair cannot be manually changed.

Spanning Tree Protocol (STP)

The Spanning Tree Protocol (STP) makes your network more resilient to link failure and also provides protection from loops — one of the major causes of broadcast storms.

STP is disabled by default and has to be manually enabled on each Switch. RSTP is enabled by default.

To be fully effective, STP/RSTP must be enabled on all Switches in your network.

RSTP provides the same functionality as STP. All STP explanatory text is applicable to RSTP. For details on how the two systems differ, see "Rapid Spanning Tree Protocol (RSTP)" *on* page 40 *and* "How RSTP Differs to STP" *on* page 45.

The following sections explain more about STP and the protocol features supported by your Switch. They cover the following topics:

- What is STP?
- How STP Works
- Using STP on a Network with Multiple VLANs

The protocol is a part of the IEEE Std 802.1D, 1998 Edition bridge specification. To explain STP more effectively, your Switch will be referred to as a bridge.

Rapid Spanning Tree Protocol (RSTP)

The Rapid Spanning Tree Protocol (RSTP) is an enhanced version of the Spanning Tree Protocol. RSTP implements the Spanning Tree Algorithm and Protocol, as defined in the IEEE Std 802.1w-2001.

Some of the benefits of RSTP are:

- Faster determination of the Active Spanning Tree topology throughout a bridged network.
- Support for bridges with more than 256 ports.
- Support for the Fast-Forwarding configuration of edge ports provided by the 'Fast Start' feature. Fast Start allows a port that is connected to an endstation to begin forwarding traffic after only 4 seconds. During this 4 seconds RSTP (or STP) will detect any misconfiguration that may cause a temporary loop and react accordingly.

  If you have Fast Start disabled on a port, the Switch will wait for 30 seconds before RSTP (or STP) lets the port forward traffic.
- Easy deployment throughout a legacy network, through backward compatibility:
  - it will default to sending 802.1D style Bridge Protocol Data Units (BPDU's) on a port if it receives packets of this format.
  - it is possible for some ports on a Switch to operate in RSTP (802.1w) mode, and other ports, for example those connected to a legacy Switch, to operate in STP (802.1D) mode.
  - you have an option to force your Switch to use the legacy 802.1D version of Spanning Tree, if required.

What is STP?

STP is a bridge-based system that allows you to implement parallel paths for network traffic and uses a loop-detection process to:

- Find and disable the less efficient paths (that is, the paths that have a lower bandwidth).
- Enable one of the less efficient paths if the most efficient path fails.

RSTP provides the same functionality as STP. All STP explanatory text is applicable to RSTP. For details on how the two systems differ, see "How RSTP Differs to STP" *on* page 45.

As an example, Figure 8 shows a network containing three LAN segments separated by three bridges. With this configuration, each segment can communicate with the others using two paths. Without STP enabled, this configuration creates loops that cause the network to overload.

Figure 8 A network configuration that creates loops

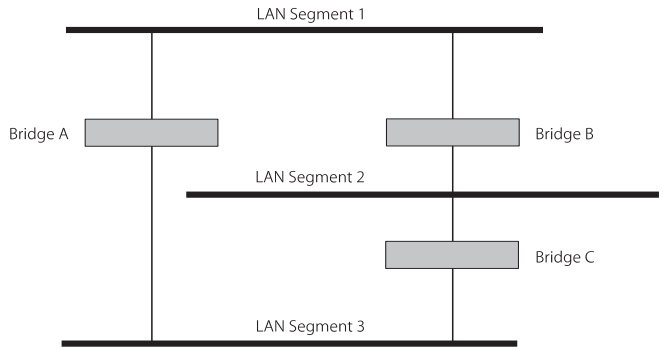
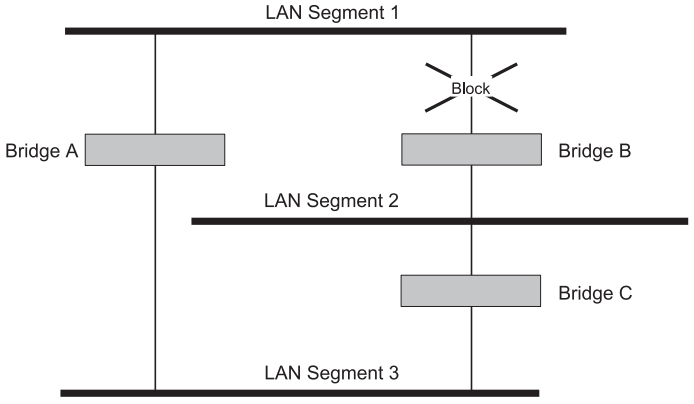


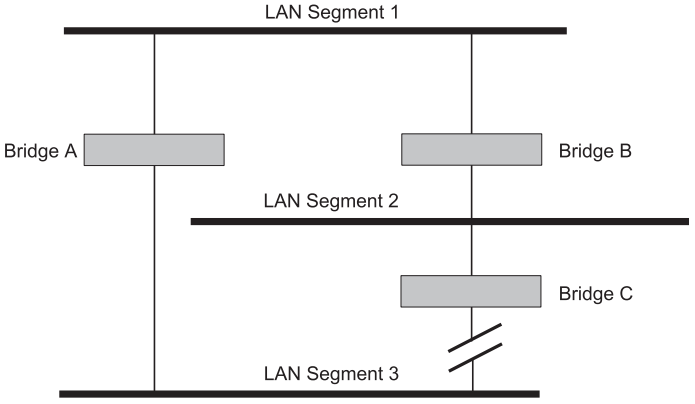
Figure 9 shows the result of enabling STP on the bridges in the configuration. STP detects the duplicate paths and prevents, or *blocks*, one of them from forwarding traffic, so this configuration will work satisfactorily. STP has determined that traffic from LAN segment 2 to LAN segment 1 can only flow through Bridges C and A, because, for example, this path has a greater bandwidth and is therefore more efficient.

Figure 9 Traffic flowing through Bridges C and A



If a link failure is detected, as shown in Figure 10, the STP process reconfigures the network so that traffic from LAN segment 2 flows through Bridge B.

Figure 10 Traffic flowing through Bridge B



STP determines which is the most efficient path between each bridged segment and a specifically assigned reference point on the network. Once the most efficient path has been determined, all other paths are blocked. Therefore, in Figure 8, Figure 9, and Figure 10, STP initially determined that the path through Bridge C was the most efficient, and so blocked the path through Bridge B. After the failure of Bridge C, STP re-evaluated the situation and opened the path through Bridge B.

How STP Works

When enabled, STP determines the most appropriate path for traffic through a network. It does this as outlined in the sections below.

STP Requirements

Before it can configure the network, the STP system requires:

- Communication between all the bridges. This communication is carried out using Bridge Protocol Data Units (BPDUs), which are transmitted in packets with a known multicast address.
- Each bridge to have a Bridge Identifier. This specifies which bridge acts as the central reference point, or Root Bridge, for the STP system — the lower the Bridge Identifier, the more likely the bridge is to become the Root Bridge. The Bridge Identifier is calculated using the MAC address of the bridge and a priority defined for the bridge. The default priority of your Switch is 32768.
- Each port to have a cost. This specifies the efficiency of each link, usually determined by the bandwidth of the link — the higher the cost, the less efficient the link. Table 4 shows the default port costs for a Switch.

Table 4 Default port costs

Port Speed	Link Type	Path Cost 802.1D-1998	Path Cost 802.1w
10 Mbps	Half Duplex	100	2,000,000
	Full Duplex	95	1,999,999
	Aggregated Link	90	1,000,000*
100 Mbps	Half Duplex	19	200,000
	Full Duplex	18	199,999
	Aggregated Link	15	100,000*
1000 Mbps	Full Duplex	4	20,000
	Aggregated Link	3	10,000*

* This path cost is correct where there are two ports in an aggregated link. However, if there are more ports in the aggregated link, the path cost will be proportionately lower. For example, if there are four ports in the aggregated link, the 802.1w path costs will be: 500,000 for 10 Mbps, 50,000 for 100 Mbps, and 5,000 for 1000 Mbps. The 802.1D-1998 path cost values are not affected by the number of ports in an aggregated link.

STP Calculation

The first stage in the STP process is the calculation stage. During this stage, each bridge on the network transmits BPDUs that allow the system to work out:

- The identity of the bridge that is to be the Root Bridge. The Root Bridge is the central reference point from which the network is configured.
- The Root Path Costs for each bridge — that is, the cost of the paths from each bridge to the Root Bridge.
- The identity of the port on each bridge that is to be the Root Port. The Root Port is the one that is connected to the Root Bridge using the most efficient path, that is, the one that has the lowest Root Path Cost. Note that the Root Bridge does not have a Root Port.
- The identity of the bridge that is to be the Designated Bridge of each LAN segment. The Designated Bridge is the one that has the lowest Root Path Cost from that segment. Note that if several bridges have the same Root Path Cost, the one with the lowest Bridge Identifier becomes the Designated Bridge.

  All traffic destined to pass in the direction of the Root Bridge flows through the Designated Bridge. The port on this bridge that connects to the segment is called the Designated Bridge Port.

STP Configuration

After all the bridges on the network have agreed on the identity of the Root Bridge, and have established the other relevant parameters, each bridge is configured to forward traffic only between its Root Port and the Designated Bridge Ports for the respective network segments. All other ports are blocked, which means that they are prevented from receiving or forwarding traffic.

STP Reconfiguration

Once the network topology is stable, all the bridges listen for Hello BPDUs transmitted from the Root Bridge at regular intervals. If a bridge does not receive a Hello BPDU after a certain interval (the Max Age time), the bridge assumes that the Root Bridge, or a link between itself and the Root Bridge, has gone down. The bridge then reconfigures the network to cater for the change. If you have configured an SNMP trap destination, when the topology of your network changes, the first bridge to detect the change sends out an SNMP trap.

CAUTION: *Network loops can occur if aggregated links are manually configured incorrectly, that is, the physical connections do not match the assignment of ports to an aggregated link. RSTP and STP may not detect these loops. So that RSTP and STP can detect all network loops you must ensure that all aggregated links are configured correctly.*

How RSTP Differs to STP

RSTP works in a similar way to STP, but it includes additional information in the BPDUs. This information allows each bridge to confirm that it has taken action to prevent loops from forming when it wants to enable a link to a neighboring bridge. This allows adjacent bridges connected via point-to-point links to enable a link without having to wait to ensure all other bridges in the network have had time to react to the change.

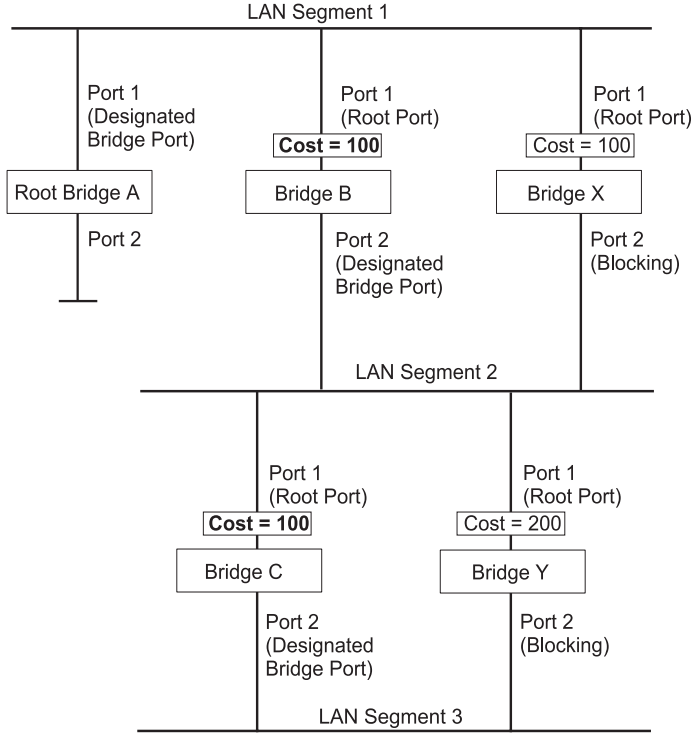
So the main benefit of RSTP is that the configuration decision is made locally rather than network-wide which is why RSTP can carry out automatic configuration and restore a link faster than STP.

STP is disabled by default and has to be manually enabled on each Switch. RSTP is enabled by default.

STP Example

Figure 11 shows a LAN that has STP enabled. The LAN has three segments, and each segment is connected using two possible links.

Figure 11 Port costs in a network



- Bridge A has the lowest Bridge Identifier in the network, and has therefore been selected as the Root Bridge.
- Because Bridge A is the Root Bridge, it is also the Designated Bridge for LAN segment 1. Port 1 on Bridge A is therefore selected as the Designated Bridge Port for LAN Segment 1.
- Port 1 of Bridges B, C, X and Y have been defined as Root Ports because they are the nearest to the Root Bridge and therefore have the most efficient path.
- Bridges B and X offer the same Root Path Cost for LAN segment 2, however, Bridge B has been selected as the Designated Bridge for the segment because it has a lower Bridge Identifier. Port 2 on Bridge B is therefore selected as the Designated Bridge Port for LAN Segment 2.
- Bridge C has been selected as the Designated Bridge for LAN segment 3, because it offers the lowest Root Path Cost for LAN Segment 3:
  - the route through Bridges C and B costs 200 (C to B=100, B to A=100)
  - the route through Bridges Y and B costs 300 (Y to B=200, B to A=100).

  Port 2 on Bridge C is therefore selected as the Designated Bridge Port for LAN Segment 3.

STP Configurations

Figure 12 shows three possible STP configurations.

- **Configuration 1 — Redundancy for Backbone Link**

  In this configuration, the Switches both have STP enabled and are connected by two links. STP discovers a duplicate path and blocks one of the links. If the enabled link breaks, the disabled link becomes re-enabled, therefore maintaining connectivity.

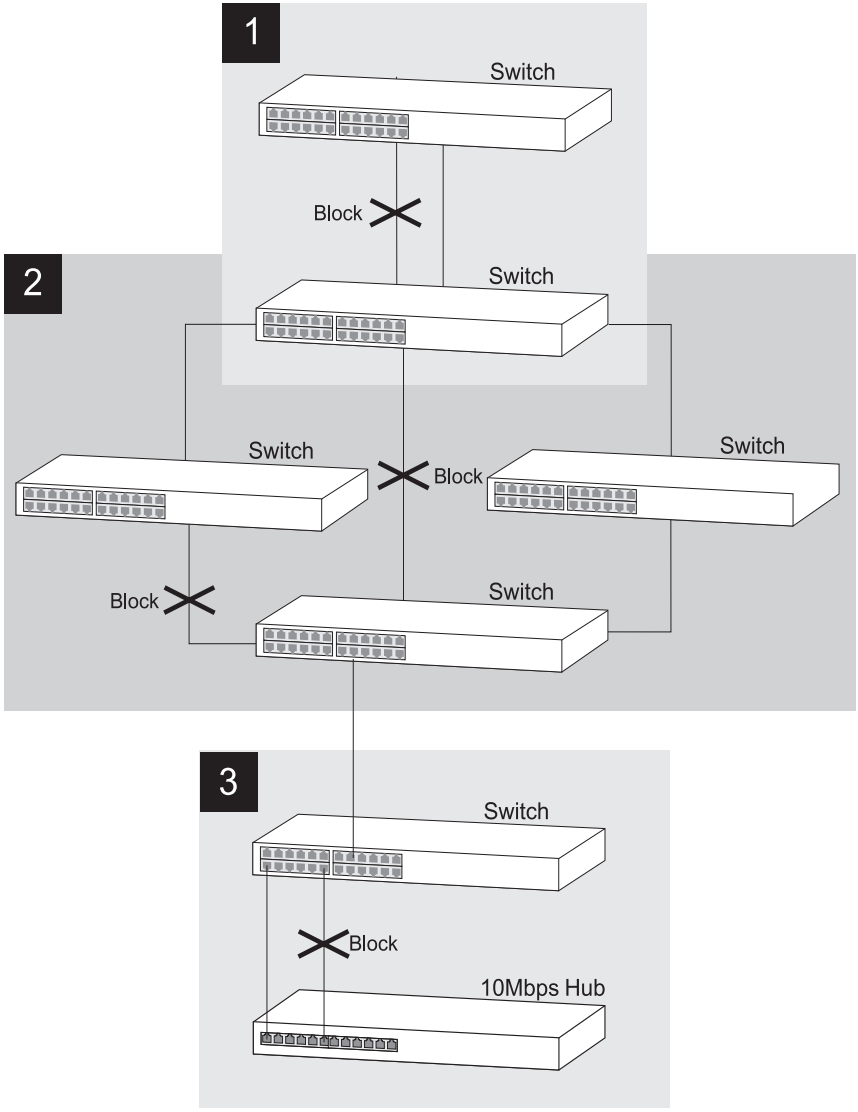
- **Configuration 2 — Redundancy through Meshed Backbone**

  In this configuration, four Switch units are connected in a way that creates multiple paths between each one. STP discovers the duplicate paths and blocks two of the links. If an enabled link breaks, one of the disabled links becomes re-enabled, therefore maintaining connectivity.

- **Configuration 3 — Redundancy for Cabling Error**

  In this configuration, a Switch has STP enabled and is accidentally connected to a hub using two links. STP discovers a duplicate path and blocks one of the links, therefore avoiding a loop.

Figure 12 STP configurations



Default Behavior

This section contains important information to note when using the RSTP and Fast Start features, particularly if you already have existing Switch 4900 units in your network with an older version of software.

RSTP Default Behavior

When using the RSTP feature on version 2.5 or later software, note the following:

- A Switch with version 2.5 (or later) software factory loaded will have RSTP enabled by default. (A Switch with version 1.x software will have STP disabled by default.)
- A Switch that you upgrade to version 2.5 (or later) software will retain its settings from prior to the upgrade, for example, if STP is disabled prior to the upgrade, it will stay disabled even though version 2.5 (or later) has RSTP enabled by default. However, if you initialize an upgraded Switch, this will clear the settings and the Switch will then assume all the version 2.5 (or later) default settings, including RSTP enabled.
- If you connect a new Switch with version 2.5 (or later) already loaded to a Distributed Fabric of upgraded units, all the upgraded units will assume the default settings of the new Switch, that is, they will have RSTP enabled by default.

Fast Start Default Behavior

When using the Fast Start feature on version 2.5 or later software, note the following:

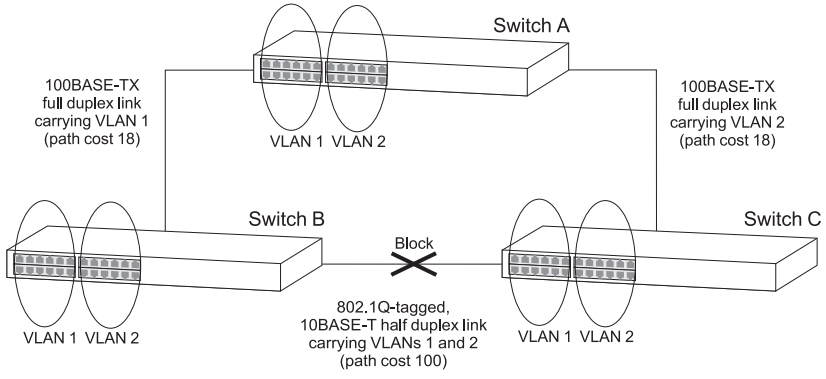
- A Switch with version 2.5 (or later) software factory loaded will have Fast Start enabled by default on the front panel ports, and disabled on any expansion module ports. (A Switch with version 1.x software will have Fast Start disabled by default.)
- A Switch that you upgrade to version 2.5 (or later) software will retain its settings from prior to the upgrade *only* if any manual settings were configured. However, if the Switch was still operating in its default state, then upon upgrade it will assume version 2.5 (or later) Fast Start default settings.
- If you initialize an upgraded Switch, this will clear the settings and the Switch will assume all the default version 2.5 (or later) settings, that is, it will have Fast Start enabled.

Using STP on a Network with Multiple VLANs

The IEEE Std 802.1D, 1998 Edition does not take into account VLANs when it calculates STP information — the calculations are only performed on the basis of physical connections. For this reason, some network configurations can result in VLANs being subdivided into a number of isolated sections by the STP system. Therefore, you must ensure that any VLAN configuration on your network takes into account the expected STP topology and alternative topologies that may result from link failures.

For example, Figure 13 shows a network containing VLANs 1 and 2. They are connected using the 802.1Q-tagged link between Switch B and Switch C. By default, this link has a path cost of 100 and is automatically blocked because the other Switch-to-Switch connections have a path cost of 36 (18+18). This means that both VLANs are now subdivided — VLAN 1 on Switch units A and B cannot communicate with VLAN 1 on Switch C, and VLAN 2 on Switch units A and C cannot communicate with VLAN 2 on Switch B.

Figure 13 Configuration that separates VLANs



To avoid any VLAN subdivision, it is recommended that all inter-Switch connections are made members of all available IEEE 802.1Q-1998 standard VLANs to ensure connectivity at all times. For example, the connections between Switches A and B, and between Switches A and C should be 802.1Q tagged and carrying VLANs 1 and 2 to ensure connectivity.

For more information about VLAN Tagging, see Chapter 8 "Setting Up Virtual LANs".

5 USING THE SWITCH DATABASE

What is the Switch Database?

The Switch Database is used by the Switch to determine where a packet should be forwarded to, and which port should transmit the packet if it is to be forwarded.

The database contains a list of entries — each entry contains three items:

- MAC (Ethernet) address information of the endstation that sends packets to the Switch.
- Port identifier, that is the port attached to the endstation that is sending the packet.
- VLAN ID of the VLAN to which the endstation belongs.

For details of the number of addresses supported by your Switch database, please refer to the Management Quick Reference Guide that accompanies your Switch.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM supplied with your Switch or on the 3Com Web site.

How Switch Database Entries Get Added

Entries are added to the Switch Database in one of two ways:

- The Switch can learn entries. The Switch updates its database with the source MAC address of the endstation that sent the packet, the VLAN ID, and the port identifier on which the packet is received.
- You can enter and update entries using the management interface via the **`bridge addressDatabase`** CLI command or an SNMP Network Manager.

Switch Database Entry States

Databases entries can have three states:

- *Learned* — The Switch has placed the entry into the Switch Database when a packet was received from an endstation. Note that:
  - Learned entries are removed (aged out) from the Switch Database if the Switch does not receive further packets from that endstation within a certain period of time (the *aging time*). This prevents the Switch Database from becoming full with obsolete entries by ensuring that when an endstation is removed from the network, its entry is also removed from the database.
  - Learned entries are removed from the Switch Database if the Switch is reset or powered-down.
- *Non-aging learned* — If the aging time is set to 0 seconds, all learned entries in the Switch Database become non-aging learned entries. This means that they are not aged out, but they are still removed from the database if the Switch is reset or powered-down.
- *Permanent* — The entry has been placed into the Switch Database using the management interface. Permanent entries are not removed from the Switch Database unless they are removed using the Switch management interface via the **`bridge addressDatabase remove`** CLI command or the Switch is initialized.

6 USING TRAFFIC PRIORITIZATION

Using the traffic prioritization capabilities of your Switch provides Quality of Service (QoS) to your network through increased reliability of data delivery. You can prioritize traffic on your network to ensure that high priority data is transmitted with minimum delay.

Traffic can be controlled by a set of rules to obtain the required Quality of Service for your network. The rules define different types of traffic and specify how each type should be treated as it passes through the Switch, for example, prioritized or discarded. Being able to define exactly how you want your Switch to treat selected applications, devices, users and types of traffic allows you to have more control over your network.

There are two different categories of rules:

- **Application-based rules** — describe how to deal with traffic for a specific application, for example, Netmeeting or Lotus Notes.
- **Device-based rules** — describe how to deal with traffic that flows to and from specific devices, for example, servers or server farms.

This chapter explains more about traffic prioritization.

For a list of the features supported by your Switch, please refer to the Management Quick Reference Guide that accompanies your Switch.

For detailed descriptions of the Command Line Interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM that accompanies your Switch.

What is Traffic Prioritization?

Today's application traffic consists of three common types of data:

- Time critical data such as video and voice.
- Business critical data such as database transactions and online transactions.
- Opportunistic data such as web browsing, email and file transfers.

When these different types of data compete for the same bandwidth, a network can quickly become overloaded, resulting in slow response times (long latency), and application time-outs. Traffic prioritization is a mechanism that allows you to prioritize data so that time-sensitive and system-critical data can be transferred smoothly and with minimal delay over a network.

The benefits of using traffic prioritization are:

- You can control a wide variety of traffic and manage congestion on your network, therefore improving performance.
- You can assign priorities to traffic, for example, set higher priorities to time-critical or business-critical applications.
- You can provide predictable throughput for multimedia applications such as video conferencing or voice over IP platforms like the 3Com NBX, as well as minimizing traffic delay and jitter.
- You can improve network performance as the amount of traffic grows, which also reduces the need to constantly add bandwidth to the network, therefore saving cost.
- You can apply security policies through traffic filtering.

How Traffic Prioritization Works

Traffic prioritization uses the four traffic queues that are present in your Switch to ensure that high priority traffic is forwarded on a different queue from lower priority traffic. This is what provides Quality of Service (QoS) to your network.

Traffic prioritization in your Switch may be applied dependent upon two factors:

- **The level of service requested by an end-station** — the transmitting end-station sets the priority of each stream of traffic. Received traffic at the Switch is forwarded through the appropriate queue depending on its priority level for onward transmission across the network.

  or

- **The level of service configured at the Switch for incoming traffic** — the network administrator configures the Switch to prioritize or discard traffic from applications or devices. For example, converged network applications such as voice or video conferencing or business critical software such as Oracle may require a high level of service from the network.

A QoS network can differentiate between time critical data, business critical data and opportunistic data (such as email, File Transfer Protocol (FTP) and Web traffic). A QoS network also has the ability to stop unauthorized usage of the network, such as online gaming.

To achieve this level of intelligence, a QoS network incorporates five processes:

- **Traffic Classification** — a QoS network examines the traffic to identify which application or device generated the traffic.
- **Traffic Marking** — after traffic is identified, it is Marked so that other network devices can identify the data and give it the correct level of service.
- **Traffic Remarking** — if a traffic packet enters the Switch with a priority marking requesting an unacceptable level of service, the Switch can Re-mark it with a different priority value to downgrade its level of service.
- **Traffic Prioritization** — once the network can differentiate types of traffic, for example, a telephone conversation from Web surfing,

prioritization can ensure that a large download from the Internet does not disrupt the telephone conversation.

- **Dropped Traffic** — traffic can be discarded either because it has an unacceptable marking or if it is of a type that is prohibited on the network, for example, an unwanted application or to/from a prohibited device.

Traffic Classification

To determine the service level to be applied to each incoming traffic type, each packet or frame must first be classified. Traffic classification is the means of identifying which application, device or user generated the traffic.

The Switch employs several methods of classifying (identifying) traffic. These can be based on any combination of fields in the first 64 bytes of the packet, and at different levels of the 7 layer OSI model as shown in Table 5.

Table 5 Attributes on which incoming traffic can be classified (identified)

OSI Layer and Protocols	Summary of Protocols
Layer 2 ■ IEEE 802.1D priority ■ EtherType	Chatty protocols such as AppleTalk and IPX, used by a small number of older devices, can cause traffic delays. Identifying and prioritizing data based on these protocols can reduce delays. AppleTalk can be identified by its EtherType of 0x809B, and IPX can be identified by EtherType 0x8137.
Layer 3 ■ Destination IP address ■ Source IP address ■ IP protocols: (ICMP, IGMP, RSVP, etc)	Many applications are identified by their Source IP address, or IP protocol. Because servers are sometimes dedicated to single applications, such as email, the Source IP address or protocol in a packet can identify which application generated the packet.
■ DiffServ code point (DSCP)	As well as being a traffic marking mechanism, the DSCP field in the IP header can also be used to classify traffic.
Layer 4 ■ UDP / TCP Source and Destination ports for IP applications	Many applications use certain TCP or UDP sockets to communicate. By examining the socket number in the IP packet, the intelligent network can determine what type of application generated the packet. This is also known as Layer 4 switching.

Traffic Marking

After traffic has been identified through classification, it must be Marked to ensure that other devices such as Layer 2 switches or routers on the network know how to prioritize the application, device or user that generated it. The Switch uses two of the industry-standard methods of marking network traffic:

- **IEEE 802.1D** — a layer 2 marking scheme.
- **Differentiated Services (DiffServ)** — a layer 3 marking scheme.

IEEE 802.1D Traffic Marking

The IEEE Std 802.1D, 1998 Edition marking scheme is an enhancement to the IEEE Std 802.1D to enable Quality of Service in the LAN. Traffic service levels are defined in the IEEE 802.1Q 4-byte tag, which is used to carry VLAN identification as well as IEEE 802.1p priority information. The 4 byte tag immediately follows the destination MAC address and Source MAC address.

The IEEE Std 802.1D, 1998 Edition priority marking scheme assigns each frame with an IEEE 802.1p priority level between 0 and 7, which determines the level of service that type of traffic should receive. Refer to Table 6 for an example of how different traffic types can be mapped to the eight IEEE 802.1p priority levels.

Table 6 IEEE recommendation for mapping 802.1p priority levels to 802.1D traffic types

IEEE 802.1p Priority Level	IEEE 802.1D Traffic Type
0	Best Effort (default)
1	Background
2	Standard (spare)
3	Excellent Effort (business critical)
4	Controlled Load (streaming multimedia)
5	Video (interactive media), less than 100 milliseconds latency and jitter
6	Voice (interactive voice), less than 10 milliseconds latency and jitter
7	Network Control Reserved traffic

The traffic marking and prioritization supported by the Switch using layer 2 information is compatible with the relevant sections of the IEEE Std 802.1D, 1998 Edition (incorporating IEEE 802.1p).

The IEEE 802.1D standard is the most widely used prioritization scheme in the LAN environment, but it does however have some restrictions:

- It requires an additional 4-byte tag in the frame, which is normally optional in Ethernet networks. Without this tag, the scheme cannot work.
- The tag is part of the IEEE 802.1Q header, so to implement QoS at layer 2, the entire network has to implement IEEE 802.1Q VLAN tagging.
- It is only supported on a LAN and not across routed WAN links, because the IEEE 802.1Q tags are removed when the packets pass through a router.

Differentiated Services (DiffServ) Traffic Marking

DiffServ is a Layer 3 marking scheme that uses the DiffServ Code Point (DSCP) field in the IP header to store the packet priority information. DSCP is an advanced intelligent method of traffic marking because you can choose how your network prioritizes different types of traffic. DSCP uses 64 values that map to user-defined service levels, allowing you to establish more control over network traffic.

Advantages of DiffServ over IEEE 802.1D are:

- You can configure how you want your Switch to treat selected applications and types of traffic, by assigning various grades of network service to them.
- No extra tags are required in the packet (that means there is no need for VLAN tagging).
- DSCP uses the IP header of a packet and therefore priority is preserved across the Internet.
- DSCP is backward compatible with IPV4 TOS, which allows operation with any existing devices with layer 3 TOS enabled prioritization scheme in use.

Traffic Re-Marking

Traffic entering the Switch may get downgraded or discarded depending on the network policies and Service Level Agreements (SLA). If for example a traffic packet enters the Switch with a priority marking higher than the network SLA, the rules set up by the network administrator can either be to Re-Mark the packet with a different 802.1D priority or new DSCP value, or alternatively to discard the traffic.

Traffic Prioritization

Your Switch supports Basic and Advanced Quality of Service (QoS) traffic prioritization. Basic traffic prioritization classifies traffic based on layer 2 of the OSI 7 layer model, and the Switch will prioritize the received traffic according to the priority information defined in the received packet. Advanced traffic prioritization can classify traffic at layers 2, 3 and 4 of the OSI 7 layer model, and treat traffic according to the rules set up by the network administrator.

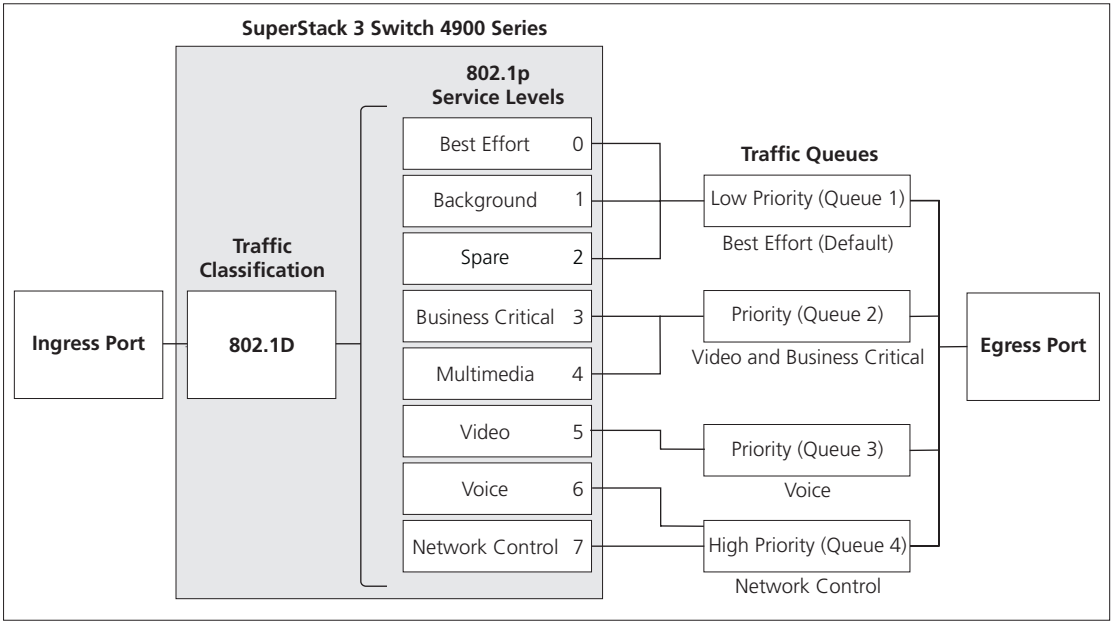
Basic Traffic Prioritization

Incoming traffic is classified based upon the IEEE 802.1D frame and is assigned to the appropriate priority queue based upon the IEEE 802.1p service level value defined in that packet. Service level markings (values) are defined in the IEEE 802.1Q 4-byte tag, and therefore traffic will only contain 802.1p priority markings if the network is configured with VLANs and VLAN tagging. All SuperStack 3 Switch 4900 Series units support basic traffic prioritization. The traffic flow through the Switch is as follows:

1. A packet received by the Switch may or may not have an 802.1p tag associated with it. If it does not, then it is given a default 802.1p tag (which is usually 0). A packet may also be discarded by the Switch in which case the packet would go no further. Alternatively, the packet may be remarked with a new 802.1p value, which will result in all knowledge of the old 802.1p tag being lost.
2. Because the 802.1p priority levels are fixed to the traffic queues (as shown in Figure 14 on page 60), the packet will be placed in the appropriate priority queue, ready for transmission through the appropriate egress port(s). When the packet reaches the head of its queue and is about to be transmitted the device determines whether or not the egress port is tagged for that VLAN. If it is, then the new 802.1p tag is used in the extended 802.1D header. If the packet is untagged it will be transmitted untagged. The new 802.1p piority cannot be inserted into an untagged packet.

The IEEE 802.1D standard specifies eight distinct levels of priority (0 to 7), each of which relates to a particular type of traffic. The priority levels and their traffic types are shown in Figure 14 in order of increasing priority. The mapping from 802.1p level to traffic queue in the Switch is proprietary and is slightly different to the recommended IEEE mapping.

Figure 14 IEEE 802.1p priority levels and recommended IEEE 802.1D traffic types



The number of queues and their mappings to the 8 levels is proprietary and can even vary between Switches from the same vendor.

You cannot alter the mapping between the IEEE 802.1p priorities and the traffic queues. These are calculated to be the most efficient, and are fixed as illustrated in Figure 14.

Figure 14 shows how traffic prioritization works at layer 2. The Switch will check a packet received at the ingress port for IEEE 802.1D traffic classification, and then prioritize it based upon the IEEE 802.1p value (service levels) in that tag. It is this 802.1p value that determines which traffic queue the packet is mapped to.

Advanced Traffic Prioritization

Incoming traffic can be classified based on packet attributes at different layers of the OSI 7 layer model. More importantly however, the network administrator has more control and can define what service level to apply to each classified traffic type, and therefore how it is treated by the Switch once it has been identified.The Switch can look in the packet for layer 2, 3 and 4 attributes to identify incoming traffic.

Most of the current applications, for example Microsoft Word, Lotus Notes and NetMeeting, are not QoS-aware and do not apply a service level to the traffic that they send. Being an intelligent Switch, your Switch can use its own rules to classify and mark the traffic. If the incoming traffic has pre-defined service level markings, however, the advanced traffic prioritization of your Switch allows you to modify and assign the appropriate DSCP and 802.1D service level markings to that incoming traffic.

The advanced traffic prioritization in the Switch allows you to:

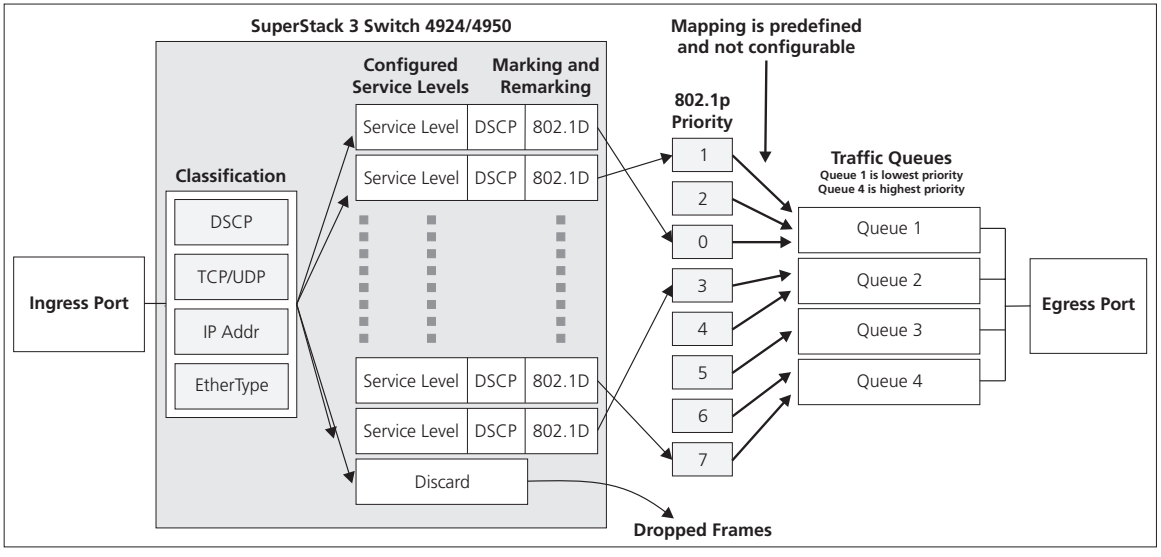
- Classify traffic based on different packet attributes. The four common methods of classification are DSCP, TCP/UDP ports, IP Address* and EtherType*.
- Mark traffic as it enters the Switch with the appropriate DSCP* and 802.1D* markings.

*Not supported by the Switch 4900 or Switch 4900 SX.

Figure 15 shows how traffic is treated using the advanced traffic prioritization in the Switch.

The DSCP field in the IP packet header can be used to classify (identify) traffic as well as carrying the priority markings, as shown in Figure 15.

Figure 15 Advanced traffic prioritization and marking



The procedure below describes how advanced traffic prioritization and marking/remarking works on the SuperStack 3 Switch 4924 and 4950. For a summary of the differences on how the feature operates on the Switches 4900 and 4900 SX, refer to Table 7.

1. The packet received at the ingress port is checked for any of the supported traffic classification methods (DSCP, TCP/UDP ports, IP Address, EtherType) to identify the traffic.
2. The classification in an incoming packet will be compared with the predefined classifications in the Switch, and if there is a match, the configured service level associated with the classified traffic will be applied.
3. The service level associated with the classifier may cause the 802.1p tag to be remarked, if the packet already has an 802.1p tag, and the DSCP value in IP packets to be remarked, or it may cause the Switch to drop the packet.
4. Remarking the 802.1p tag, DSCP field or dropping the packet are optional, and have to be configured by the network administrator.
5. It is the priority associated with the packet that is used to direct it to the appropriate queue. This is determined as follows:
   - If the packet matches a classifier with a configured service level specifying that the DSCP or 802.1p tag should be re-marked, then the packet is re-marked with the configured DSCP value and/or the 802.1p priority.

You can have both an 802.1p tag and a DSCP remarked. This is possible as the 802.1p tag is Layer 2 and requires VLAN tagging to be enabled, and DSCP is Layer 3 and uses the TOS field in the IP header.

   - Otherwise, if there are no other classifiers except the 802.1p tag, then the packet will pass through the Switch with the original 802.1p priority tag.
   - Otherwise, if the received packet does not have an 802.1p tag, then a default 802.1p tag (which is usually 0) is assigned to it.

Traffic Queues

It is the multiple traffic queues within the Switch hardware that allow packet prioritization to occur. Higher priority traffic can pass through the Switch without being delayed by lower priority traffic. As each packet arrives in the Switch, it passes through any ingress processing (which includes classification, marking/remarking or dropping), and is then

sorted into the appropriate queue. The Switch then forwards packets from each queue. It is worth noting that each egress port has its own set of queues, so that if one port is congested it does not interfere with the queue operation of other ports.

The Switch uses the Weighted Round Robin (WRR) queuing mechanism. This method services all the traffic queues, giving priority to the higher priority queues. Under most circumstances, this method gives high priority precedence over low-priority, but in the event that high-priority traffic exceeds the link capacity, lower priority traffic is not blocked.

Configuring Traffic Prioritization on the Switch

Your Switch allows you to discard and prioritize applications as well as devices to obtain Quality of Service (QoS) for your network. Configuring traffic prioritization on the Switch is referred to as QoS configuration, and you must follow the following sequence of steps:

1. **Traffic Classification** — first identify the types of traffic requiring special treatment. This is called creating a Classifier. Your Switch is capable of identifying traffic from many of the various attributes of an incoming packet, as shown in Table 5 on page 56. Traffic classification can be determined by various attributes across the seven layers of the OSI model. The Switch then groups classified traffic in order to schedule them with the appropriate service level. Your Switch has five default Classifiers as shown in Table 8 on page 68.
2. **Service Levels** — you can create and modify service levels to determine the priority that will be applied to each classified traffic type. For example, the Switch can discard the traffic or Re-mark it from the DiffServ Code Point (DSCP) to an 802.1D priority to ensure the packet is prioritized correctly by other parts of the network, or it can discard the packet. Your Switch offers 6 predefined standard service levels which are shown in Table 9 on page 68.
3. **Create a QoS Profile** — the next step is to create a QoS profile. A QoS profile can be made up of several rules. A rule is what defines how a particular traffic type should be treated by your Switch, and is made up of a Classifier and service level. Creating a QoS profile therefore is to associate Classifier(s) with service levels. QoS in your Switch uses the DSCP values to create classifiers that map to service levels, which then remark to 802.1p tags. It is the 802.1p tag that determines the queue, and so the level of service for the identified traffic type. Refer to Table 10 on page 70 for some examples of rules that can be added to a QoS profile.

4 **Apply a QoS profile** — after a QoS profile has been created, it is assigned to the selected port(s). When the profile is assigned to the port(s) or to the entire distributed fabric, the QoS configuration defined in the profile will immediately become active.

- For Switches 4900/4900SX — the QoS profile is set up on a per-Switch (unit) basis and is applied to each packet received on every port of that Switch. Only one QoS profile can be applied to each Switch.
- For Switches 4924/4950 — the QoS profile is set up on a per-port basis, though the same profile can be set to multiple or all ports, and the profile is applied to each packet received on assigned port(s) of that Switch. Only one QoS profile can be applied to each port. Multiple profiles can be set to various port(s).

Methods of Configuring Traffic Prioritization

QoS can be configured on your Switch using 3Com Network Supervisor or via the Command Line Interface (CLI).

- The 3Com Network Supervisor application supplied on the CD-ROM accompanying your Switch is the main tool for configuring QoS, and 3Com recommends that you use this application for ease of use. 3Com Network Supervisor automatically configures all QoS-capable 3Com Switches in the network to prioritize the selected applications and devices, as well as generating a report of all the configurations made.
- QoS can also be configured via the Command Line Interface (CLI), which offers a greater level of configuration. For a detailed description of the commands required to configure QoS via the CLI, please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM that accompanies your Switch.

Important QoS Considerations

Before implementing QoS on your network you need to consider the following points:

- Your Switch:
  - Has a predefined Classifier for NBX traffic, which is DSCP service level 46. If the profile assigned to the port on which the NBX traffic is received has an NBX classifier in it such as the default profile does, the Switch will automatically detect NBX telephone voice traffic and prioritize accordingly. The Switch also has an NBX classifier for Ethernet Type 0x8868, which is the layer 2 NBX traffic. NBX is layer 2 out of the box and has to be configured by the user

to be layer 3, so the DSCP 46 classifier may in fact be used in fewer NBX installations than the Ethernet Type 0x8868.

  - Can map between IEEE 802.1D and DSCP to support legacy devices in the network that only support IEEE 802.1D.
  - Has four traffic queues, but it is important to note that not all Switches have the same number of priority queues.
  - Has six service levels pre-programmed by default as listed in Table 9, but up to 900 service levels can be created.
  - Has five Classifiers pre-programmed by default, as listed in Table 8. (The Switch 4900 and 4900 SX only support four Classifiers by default.)
  - Only one QoS profile can be applied to each port.

- QoS is about providing a consistent, predictable data delivery service. It should not be used as an alternative to deploying sufficient bandwidth. The recommended configuration for most networks is 10/100 Mbps switching to the desktop, Gigabit connections for servers, and non-blocking Gigabit backbones.
- QoS requires the support of every network device from end-to-end. All devices in the network should support QoS. If there is just one section in the data path that does not support QoS, it can produce bottlenecks and slowdowns, although a performance improvement will be noticed over the parts of the network that do support QoS.
- Ensure that all QoS devices are configured the same way. Mismatches will cause the same traffic to be prioritized in one section and not in another. Use a comprehensive QoS management package, such as 3Com Network Supervisor, that will configure all devices in the network simultaneously and check for errors.
- Only use Switches or hardware-based routers in the LAN. Hubs cannot prioritize traffic, and software-based routers can cause bottlenecks.
- Use Switches and hardware-based routers that understand both the IEEE 802.1D (incorporating IEEE 802.1p) and DSCP marking schemes.
- Classify traffic as soon as it enters the network. If traffic is not classified until it gets to the WAN router or firewall, end-to-end prioritization cannot be guaranteed. The ideal place for traffic classification is within the Switch.
- Traffic Marking is performed as a result of classification, and so you should aim to perform the marking only once to reduce the additional

requirements that QoS places upon the capabilities of your network infrastructure.

- As DSCP uses a field in the IP header, it is only possible to use the DSCP in IP packets. It does not apply, for example, to AppleTalk, IPX or NetBEUI.
- Because DSCP is a redefinition of the use of the TOS byte in the IP header, there are some issues with interaction with IP TOS based networks.

You need to consider the following when setting QoS profiles on your Switch.

- The Switch 4900 Series supports four queues (two physical and two logical) per port by default.
- Traffic can be classified based on 802.1D priority as set up by the end station devices.
- Physical queues are classified based on latency, logical queues are based on loss.
- Traffic prioritization is based on 802.1D tag information, DiffServe Code Point (DSCP), TCP/UDP, IP address (Switch 4924/4950), and Ethertype (Switch 4924/4950).

Refer to the following table for implementation considerations:

Table 7 Traffic Prioritization and QoS Features

Traffic Prioritization /QoS Feature	SuperStack 3 Switch 4900/4900 SX	SuperStack 3 Switch 4924/4950
Queueing		
Queues	2/2	2/2
Queue Scheduling	Weighted Round Robin (WRR)*	Weighted Round Robin (WRR)
Classification		
802.1p	Yes	Yes
TCP/UDP	Yes (4 classifiers)	Yes (6 classifiers)
DSCP	Yes (0-63 classifiers)	Yes (no zero allowed)
IP Address	No	Yes (12)
EtherType	No	Yes (1)
Service Levels		
DSCP remarking	No	Yes (per port)
Immediate discard	No	Yes
QoS profiles	Per unit	Per port

* Weighted Round Robin (WRR) is the method the Switch uses internally to service the queues. There is an upper limit on the number of packets that can be transmitted from each queue. When the limit is reached, packets from the next lower queue are then selected, and so on. This process continually repeats and ensures that packets from the lowest queues are not overlooked. This weighting is not user configurable.

Default QoS Configurations

The Switch has some pre-configured defaults which are listed in Table 8 and Table 9.

Table 8 Default traffic classifiers configured in your Switch

	Classifier Name	Classifier Type	Protocol Identifier	Used in QoS Profile
2*	3Com NBX Voice-LAN	EtherType	0x8068	2
3	3Com NBX Voice-IP	DSCP	46	2
4	Web-HTTP	IpPort	TCP (80)	None
5	Network management - SNMP	IpPort	UDP (161)	-
6	Network management - SNMP Traps	IpPort	UDP (162)	-

* Not supported by Switch 4900 and 4900 SX.

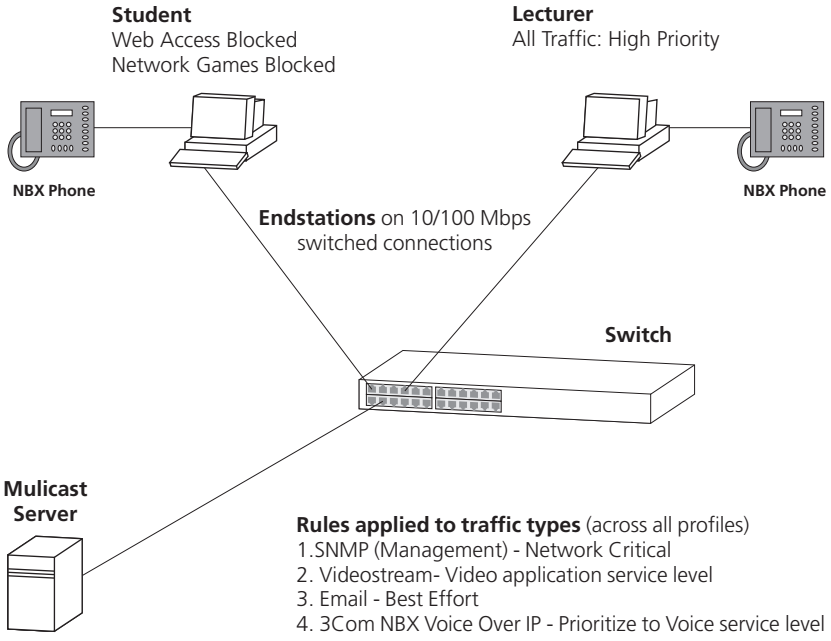
Table 9 Default service levels configured in your Switch

	Service Level Name	IEEE 802.1D Priority	DSCP Marking	Used in QoS Profile
1	Drop	-	-	None
2	Best Effort	0	0	-
3	Business Critical	3	16	None
4	Video Applications	5	24	None
5	Voice Applications	6	46	2
6	Network Control	7	48	None

Example QoS Configurations

Figure 16 shows a simple example of how QoS can be implemented on a university campus. It shows how traffic receives the appropriate prioritization and treatment across the network according to the applications used (traffic type), at which location the end user is located, and also the port type upon which the data is received. All of this is determined by the setup of the QoS profiles applied to the Switch port.

Figure 16 University campus QoS network example



See "Utilizing the Traffic Prioritization Features of Your Network" on page 154 for a further network example.

Some examples of rules that can be set up and added to a QoS profile are shown in Table 10 on page 70.

Table 10 Example QoS profile rules

Rule	Example
Priority re-mapping	Mark the 802.1D priority of each packet on a port according to its DSCP to ensure correct priority treatment by non-DiffServ devices on the network.
Endstation based	Prioritize packets (by DSCP or 802.1D) destined for a particular endstation or server in the network.
Network protocol based*	Prioritize IP traffic over IPX.
NBX phone traffic†	Prioritize over data traffic.
Layer 4 port number	Prioritize Lotus Notes traffic over web (HTTP) traffic.

* Not supported by Switch 4900 and 4900 SX.
† Supported on Switch 4900 and 4900 SX at Layer 3 only (DSCP).

7 STATUS MONITORING AND STATISTICS

This chapter contains details of the Remote Monitoring (RMON) feature that assists you with status monitoring and statistics.

For a list of the features supported by your Switch, please refer to the Management Quick Reference Guide that accompanies your Switch.

RMON

Using the RMON capabilities of a Switch allows you to improve your network efficiency and reduce the load on your network.

This section explains more about RMON. It covers the following topics:

- What is RMON?
- Benefits of RMON
- RMON and the Switch

You can only use the RMON features of the Switch if you have a management application that supports RMON, for example 3Com Network Supervisor.

What is RMON?

RMON is a system defined by the IETF (Internet Engineering Task Force) that allows you to monitor the traffic of LANs or VLANs.

RMON is an integrated part of the Switch software agent and continually collects statistics about a LAN segment or VLAN, and transfers the information to a management workstation on request or when a pre-defined threshold is crossed. The workstation does not have to be on the same network as the Switch and can manage the Switch by in-band or out-of-band connections.

The RMON Groups

The IETF define groups of Ethernet RMON statistics. This section describes seven groups supported by the Switch, and details how you can use them.

Statistics

The Statistics group provides traffic and error statistics showing packets, bytes, broadcasts, multicasts and errors on a LAN segment or VLAN.

Information from the Statistics group is used to detect changes in traffic and error patterns in critical areas of your network.

History

The History group provides historical views of network performance by taking periodic samples of the counters supplied by the Statistics group.

The group is useful for analyzing the traffic patterns and trends on a LAN segment or VLAN, and for establishing the normal operating parameters of your network.

Alarms

The Alarms group provides a mechanism for setting thresholds and sampling intervals to generate events on any RMON variable.

Alarms are used to inform you of network performance problems and they can trigger automated responses through the Events group.

Hosts

The Hosts group specifies a table of traffic and error statistics for each host (endstation) on a LAN segment or VLAN. Statistics include packets sent and received, octets sent and received, as well as broadcasts, multicasts, and error packets received.

The group supplies a list of all hosts that have transmitted across the network.

Hosts Top N

This group requires implementation of the Hosts group. The Hosts Top N group extends the Hosts table by providing sorted host statistics, such as the top 20 hosts sending packets or an ordered list of all hosts according to the errors they sent over the last 24 hours.

Matrix

The Matrix group shows the amount of traffic and number of errors between pairs of devices on a LAN segment or VLAN. For each pair, the Matrix group maintains counters of the number of packets, number of octets, and number of error packets between the hosts.

The conversation matrix helps you to examine network statistics in more detail to discover, for example, who is talking to whom or if a particular PC is producing more errors when communicating with its file server. Combined with Hosts Top N, this allows you to view the busiest hosts and their primary conversation partners.

Events

The Events group provides you with the ability to create entries in an event log and send SNMP traps to the management workstation. Events are the action that can result from an RMON alarm. In addition to the standard five traps required by SNMP (link up, link down, warm start, and cold start), RMON adds two more: rising threshold and falling threshold.

Effective use of the Events group saves you time; rather than having to watch real-time graphs for important occurrences, you can depend on the Event group for notification. Through the SNMP traps, events can trigger other actions, therefore providing a way to automatically respond to certain occurrences.

Benefits of RMON

Using the RMON features of your Switch has three main advantages:

- **It improves your efficiency**

  Using RMON allows you to remain at one workstation and collect information from widely dispersed LAN segments or VLANs. This means that the time taken to reach a problem site, set up equipment, and begin collecting information is largely eliminated.

- **It allows you to manage your network in a more proactive manner**

  If configured correctly, RMON can deliver information before problems occur. This means that you can take action before they affect users. In addition, the software records the behavior of your network, so that you can analyze the causes of problems.

- **It reduces the load on the network and the management workstation**

  Traditional network management involves a management workstation polling network devices at regular intervals to gather statistics and identify problems or trends. As network sizes and traffic levels grow, this approach places a strain on the management workstation and also generates large amounts of traffic.

  RMON, however, autonomously looks at the network on behalf of the management workstation without affecting the characteristics and performance of the network. RMON reports by exception, which means that it only informs the management workstation when the network has entered an abnormal state.

RMON and the Switch

The RMON support provided by your Switch is detailed in Table 11

Table 11 RMON support supplied by the Switch

RMON group	Support supplied by the Switch
Statistics	A new or initialized Switch has one Statistics session per port and one default Statistics session for VLAN 1.
History	A new or initialized Switch has two History sessions per port, and one default History session for VLAN 1. These sessions provide the data for the web interface history displays: ■ 30 second intervals, 120 historical samples stored ■ 2 hour intervals, 96 historical samples stored
Alarms	A new or initialized Switch has two alarms defined for each port: ■ Broadcast bandwidth used. ■ Percentage of errors over one minute You can modify these alarms using an RMON management application, but you cannot create or delete them. You can define up to 200 alarms for the Switch. For more information about the alarms setup on the Switch, see "Alarm Events" on page 75 and "The Default Alarm Settings" on page 76.
Hosts	Although Hosts is supported by the Switch, Hosts sessions are defined on VLANs only (default VLAN 1). There are no Hosts sessions defined on a new or initialized Switch.
Hosts Top N	Although Hosts Top N is supported by the Switch, there are no Hosts Top N sessions defined on a new or initialized Switch.

Table 11 RMON support supplied by the Switch

RMON group	Support supplied by the Switch
Matrix	Although Matrix is supported by the Switch, Matrix sessions are defined on VLANs only (default VLAN 1). There are no Matrix sessions defined on a new or initialized Switch.
Events	A new or initialized Switch has Events defined for use with the default alarm system. See "The Default Alarm Settings" on page 76 for more information.

When using the RMON features of the Switch, note the following:

- After the default sessions are created, they have no special status. You can delete or change them as required.
- The greater the number of RMON sessions, the greater the burden on the management resources of the Switch. If you have many RMON sessions, the forwarding performance of the Switch is not affected but you may experience slow response times from the web interface.

Alarm Events

You can define up to 200 alarms for the Switch. The events that you can define for each alarm and their resulting actions are listed in Table 12.

Table 12 Alarm Events

Event	Action
No action	
Notify only	Send Trap.
Notify and filter port	Send Trap. Block broadcast and multicast traffic on the port. Recovers with the *unfilter port* event.
Notify and disable port	Send Trap. Turn port off.
Notify and enable port	Send Trap. Turn port on.
Disable port	Turn port off.
Enable port	Turn port on.
Notify and switch resilient port	Send Trap. If port is the main port of a resilient link pair then move to standby.
Notify and unfilter port	Send Trap. Stop blocking broadcast and multicast traffic on the port.
System started	
Software Upgrade report	

The Default Alarm Settings

A new or initialized Switch has two alarms defined for each port:

- Broadcast bandwidth used.
- Percentage of errors over one minute

The default values and actions for each of these alarms are given in Table 13.

Table 13 Values for the default alarms

Statistic	High Threshold	Low Threshold Recovery	Period
Broadcast bandwidth used	Value: 20%	Value: 10%	30 secs
	Action: Notify and filter	Action: Notify and unfilter	
Number of errors over 10 seconds	Value: 8 errors per 10 seconds	Value: 8 errors per 10 seconds	10 secs
	Action: Smart auto-sensing will reduce port speed	Action: None. (Speed can only be increased upon link loss, for example by removing and replacing the cable, or by triggering the port to perform another auto-negotiation on that link.)	

The Audit Log

The Switch keeps an audit log of all management user sessions, providing a record of a variety of changes, including ones relating to RMON. The log can only be read by users at the *security* access level using an SNMP Network Management application.

Each entry in the log contains information in the following order:

- Entry number
- Timestamp
- User ID
- Item ID (including qualifier)
- New value of item

The last 16 operations are stored in the audit log. The oldest records are overwritten first.

Email Notification of Events

Your Switch allows you to receive email notification when certain RMON events occur. You can receive notification via email, SMS (Short Message Service), or pager, of the event that has occurred.

This feature uses an SMTP (Simple Mail Transfer Protocol) email client to send the notification email. The SMS and pager messages are constrained on message size so they are sent to a different email address which creates the message to be displayed and then forwards it on to the SMS or pager gateway.

You can configure the Switch to send alerts via email or pager notification if certain events occur. You can configure the email address to which you wish the notifications to be sent. However, you cannot change the factory default notification messages for event emails.

The events that can generate email notification are:

- Unit powers up.
- A link fails or returns to service — you can select specific links that you wish to receive messages for, for example, a mission-critical link to a server.
- A resilient link activates.

RMON traps continue to be sent, in addition to any email notifications you may receive.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM supplied with your Switch or on the 3Com Web site.

What is Trace Route?

Trace route is a feature that allows you to find out about the route that packets take from their local device to their remote destination. This can help when troubleshooting or monitoring your network.

You can trace the route of a packet using the TraceRoute command or operation on the Protocol>IP menu.

The trace route command displays the IP address of the originating interface and the information about routes to a destination IP device. It

also displays the round-trip time of each. The screen with the trace route results will be refreshed once every second.

The trace route will stop if:

- the hop count reaches the maximum default value of 30 (user configurable via SNMP)
- it finds the target host
- it receives five consecutive failures.

8 SETTING UP VIRTUAL LANS

Setting up Virtual LANs (VLANs) on your Switch reduces the time and effort required by many network administration tasks, and increases the efficiency of your network.

This chapter explains more about the concept of VLANs and explains how they can be implemented on your Switch. It covers the following topics:

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM supplied with your Switch or on the 3Com Web site.

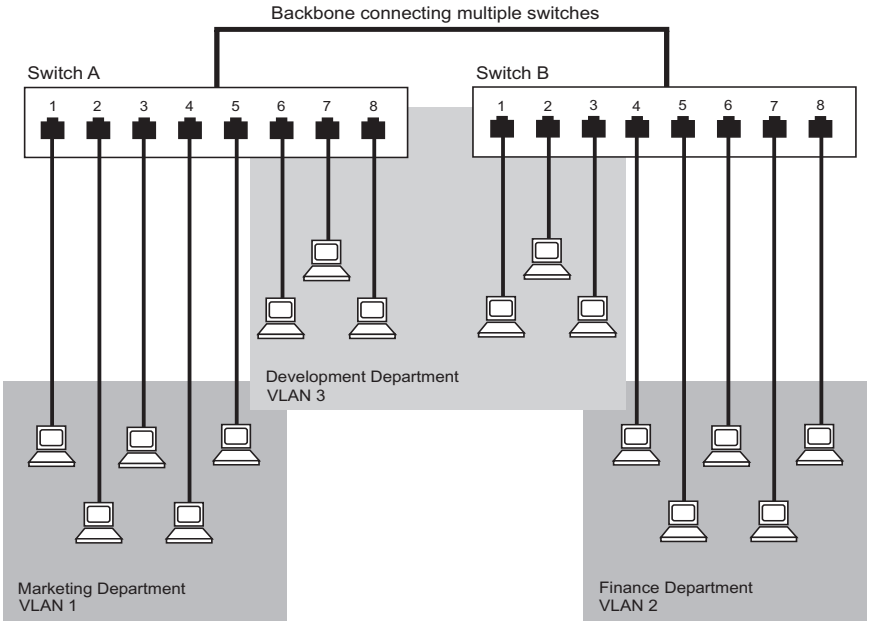
What are VLANs?

A VLAN is a flexible group of devices that can be located anywhere in a network, but which communicate as if they are on the same physical segment. With VLANs, you can segment your network without being restricted by physical connections — a limitation of traditional network design. As an example, with VLANs you can segment your network according to:

- **Departmental groups** — For example, you can have one VLAN for the Marketing department, another for the Finance department, and another for the Development department.
- **Hierarchical groups** — For example, you can have one VLAN for directors, another for managers, and another for general staff.

- **Usage groups** — For example, you can have one VLAN for users of e-mail, and another for users of multimedia.

Figure 17 A network setup showing three VLANs



Benefits of VLANs

The main benefit of VLANs is that they provide a network segmentation system that is far more flexible than any traditional network. Using VLANs also provides you with three other benefits:

- **VLANs ease the movement of devices on networks**

  With traditional networks, network administrators spend much of their time dealing with moves and changes. If users move to a different subnetwork, the addresses of each endstation must be updated manually.

  With a VLAN setup, if an endstation in VLAN *Marketing* for example is moved to a port in another part of the network, and retains its original subnet membership, you only need to specify that the new port is in VLAN *Marketing*. You do not need to carry out any re-cabling.

- **VLANs provide extra security**

  Devices within each VLAN can only communicate with other devices in the same VLAN. If a device in VLAN *Marketing* needs to communicate with devices in VLAN *Finance*, the traffic must pass through a routing device or Layer 3 Switch.

- **VLANs help to control traffic**

  With traditional networks, congestion can be caused by broadcast traffic that is directed to all network devices whether they require it or not. VLANs increase the efficiency of your network because each VLAN can be set up to contain only those devices that need to communicate with each other.

VLANs and Your Switch

Your Switch provides support for VLANs using the IEEE Std 802.1Q-1998. This standard allows traffic from multiple VLANs to be carried across one physical link.

This chapter describes VLANs as supported in a Layer 2 environment. If you are using VLANs in a Layer 3 environment, you will also need to refer to "IP Routing Concepts" *on* page 96 *which explains more about VLAN-based routing and the use of multiple IP interfaces per VLAN.*

The 802.1Q standard allows each port on your Switch to be placed in:

- Any VLAN defined on the Switch.
- Several VLANs at the same time using 802.1Q tagging.

The standard requires that you define the following information about each VLAN on your Switch before the Switch can use it to forward traffic:

- *VLAN Name* — This is a descriptive name for the VLAN (for example, Marketing or Management).
- *802.1Q VLAN ID* — This is used to identify the VLAN if you use 802.1Q tagging across your network.

If an automatic aggregated link (created by LACP) contains ports with different VLAN membership, the aggregated link will inherit the VLAN memberhip of the first port that comes up in the aggregated link. It will override any pre-defined VLAN membership for the aggregated link. You therefore need to ensure that prior to the aggregated link forming, every individual port that will be in the aggregated link has the required VLAN memberships configured.

The Default VLAN

A new or initialized Switch contains a single VLAN, the Default VLAN. This VLAN has the following definition:

- *VLAN Name* — Default VLAN
- *802.1Q VLAN ID* — 1 (if tagging required)

All the ports are initially placed in this VLAN, and it is the only VLAN that allows you to access the management software of the Switch over the network, if using a Layer 2 environment only.

If you are using VLANs in a Layer 3 environment you can access the management software of the Switch and manage it via any IP interface that you have established.

Closed VLANs

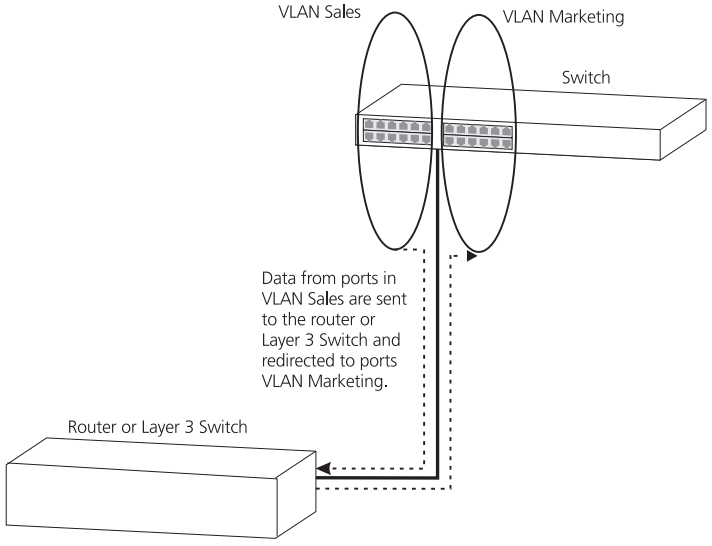
Your Switch provides "Closed VLANs". Closed VLANs enhance security because a port will only ever receive packets that are destined for that port's configured VLANs. If an IEEE 802.1Q tagged packet (that is, a packet that contains a VLAN ID) is received on a port and that port is *not* a member of that VLAN, then the packet is dropped.

Closed VLANs (also known as VLAN Ingress Filtering) are implemented on SuperStack 3 Switch 4900 Series, SuperStack 3 Switch 4400 Series, and 3Com Switches 4050 and 4060.

Communication Between VLANs

Communication between different VLANs can only take place if they are all connected to a router or a Layer 3 Switch. Alternatively, if the Switch containing the VLANs is itself a Layer 3 Switch and is configured correctly, it will be able to route the traffic from one VLAN to the other internally.

Figure 18 Two VLANs connected via a router or Layer 3 Switch



Creating New VLANs

If you want to move a port from the Default VLAN to another VLAN, you must first define information about the new VLAN on your Switch.

VLANs: Tagged and Untagged Membership

Your Switch supports 802.1Q VLAN tagging, a system that allows traffic for multiple VLANs to be carried on a single physical (backbone) link.

When setting up VLANs you need to understand when to use untagged and tagged membership of VLANs. Quite simply, if a port is in a single VLAN it can be an untagged member, but if the port needs to be a member of multiple VLANs, tagged membership must be defined. Typically endstations (for example, clients) will be untagged members of one VLAN, while inter-Switch connections will be tagged members of all VLANs.

The IEEE 802.1Q-1998 standard defines how VLANs operate within an open packet-switched network. An 802.1Q compliant packet carries additional information that allows a Switch to determine to which VLAN the port belongs. If a frame is carrying the additional information, it is known as *tagged*.

To carry multiple VLANs across a single physical (backbone) link, each packet must be tagged with a VLAN identifier so that the Switches can identify which packets belong in which VLANs. To communicate between VLANs a router must be used.

VLAN Configuration Examples

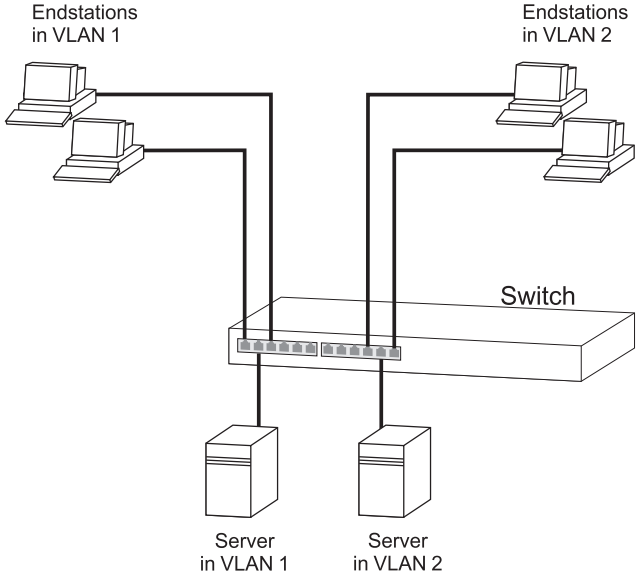
This section contains examples of VLAN configurations. It describes how to set up your Switch to support simple untagged and tagged connections.

Using Untagged Connections

The simplest VLAN operates in a small network using a single Switch. In this network there is no requirement to pass traffic for multiple VLANs across a link. All traffic is handled by the single Switch and therefore untagged connections can be used.

The example shown in Figure 19 illustrates a single Switch connected to endstations and servers using untagged connections. Ports 1, 2 and 3 of the Switch belong to VLAN 1, ports 10, 11 and 12 belong to VLAN 2. VLANs 1 and 2 are completely separate and cannot communicate with each other. This provides additional security for your network.

Figure 19 VLAN configuration example: Using untagged connections



To set up the configuration shown in Figure 19:

1 **Configure the VLANs**
Define VLAN 2 on the Switch. VLAN 1 is the default VLAN and already exists.

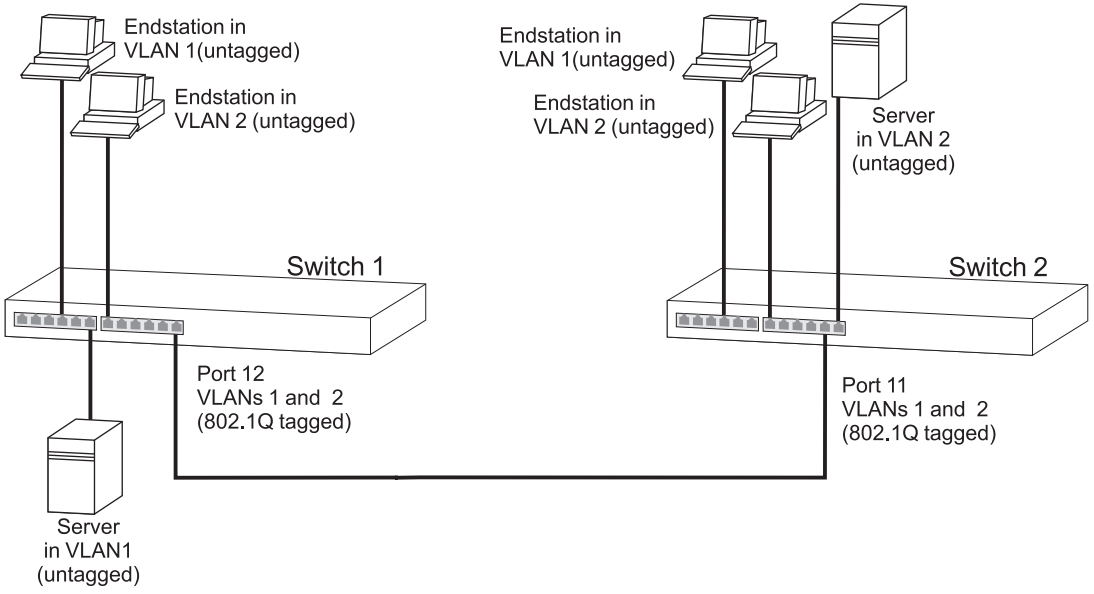
2 **Add ports to the VLANs**
Add ports 10, 11 and 12 of the Switch as untagged members to VLAN 2.

Using 802.1Q Tagged Connections

In a network where the VLANs are distributed amongst more than one Switch, you must use 802.1Q tagged connections so that all VLAN traffic can be passed along the links between the Switches. 802.1Q tagging can only be used if the devices at both ends of a link support IEEE 802.1Q.

The example shown in Figure 20 illustrates two Switch units. Each Switch has endstations and a server in VLAN 1 and VLAN 2. All endstations in VLAN 1 need to be able to connect to the server in VLAN1 which is attached to Switch 1 and all endstations in VLAN 2 need to connect to the server in VLAN2 which is attached to Switch 2.

Figure 20 VLAN configuration example: 802.1Q tagged connections



To set up the configuration shown in Figure 20:

1 **Configure the VLANs on Switch 1**
Define VLAN 2. VLAN 1 is the default VLAN and already exists.

2 **Add endstation ports on Switch 1 to the VLANs**
Place the endstation ports in the appropriate VLANs as untagged members.

3 **Add port 12 on Switch 1 to the VLANs**
Add port 12 on Switch 1 as a tagged member of both VLANs 1 and 2 so that all VLAN traffic is passed over the link to Switch 2.

4 **Configure the VLANs on Switch 2**
Define VLAN 2. VLAN 1 is the default VLAN and already exists.

5 **Add endstation ports on Switch 2 to the VLANs**
Place the endstation ports in the appropriate VLANs as untagged members.

6 **Add port 11 on Switch 2 to the VLANs**
Add port 11 on Switch 2 as a tagged member of both VLANs 1 and 2 so that all VLAN traffic is passed over the link to Switch 1.

7 **Check the VLAN membership for both Switches**
The relevant ports should be listed in the VLAN members summary.

8 **Connect the Switches**
Connect port 12 on Switch 1 to port 11 on Switch 2.

The VLANs are now configured and operational and the endstations in both VLANs can communicate with their relevant servers.

9 USING AUTOMATIC IP CONFIGURATION

This chapter explains more about IP addresses and how the automatic IP configuration option works. It covers the following topics:

- How Your Switch Obtains IP Information
- How Automatic IP Configuration Works
- Important Considerations

For detailed information on setting up your Switch for management, see the Getting Started Guide that accompanies your Switch.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM supplied with your Switch or on the 3Com Web site.

For background information on IP addressing, see Appendix C "IP Addressing".

How Your Switch Obtains IP Information

You can use one of the following methods to allocate IP information to your Switch (essential if you wish to manage your Switch across the network):

- **Automatic IP Configuration** (default) — the Switch tries to configure itself with IP information. It uses the following industry standard methods to automatically allocate the Switch IP information:
  - Dynamic Host Configuration Protocol (DHCP)
  - Auto-IP — the Switch will configure itself with its default IP address 169.254.100.100 if it is operating in a standalone mode, and/or no other Switches on the network have this IP address. If this default IP address is already in use on the network then the Switch detects this and configures itself with an IP address in the range 169.254.1.0 to 169.254.254.255.
  - Bootstrap Protocol (BOOTP)

  For ease of use, you do not have to choose between these three automatic configuration methods. The Switch tries each method in a specified order as described in "Automatic Process" on page 89.
- **Manual IP Configuration** — you can manually input the IP information (IP address, subnet mask, and default gateway).

If you select `none` *for no IP configuration the Switch will not be accessible from a remote management workstation on the LAN. In addition, the Switch will not be able to respond to SNMP requests.*

How Automatic IP Configuration Works

When your Switch is powered up for the first time the IP configuration setting is set to `automatic` — this is the default setting.

If your Switch has been powered up before, whichever of the three options for IP configuration (`manual, automatic, none`) was last configured is activated when the Switch powers up again.

You can switch to manual IP configuration at any time using a serial port or known IP address connection to set up the IP information. For more information see the Getting Started Guide that accompanies your Switch.

Automatic Process To detect its IP information using the automatic configuration process, the Switch goes through the following sequence of steps:

1 The DHCP client that resides in the Switch makes up to four attempts to contact a DHCP server on the network requesting IP information from the server. The attempts are at 0, 4, 12, 28 second intervals.

- If a DHCP server is on the network and working correctly it responds to the clients request with an IP address (allocated from a pool of available addresses) and other parameters such as a subnet mask, default gateway, lease time, and any other options configured in the DHCP server.

The way a DHCP server responds is dependant on the DHCP server settings. Therefore the way your DHCP server responds may be different to the process outlined.

- If the DHCP process fails after 30 seconds on all four attempts, then the Switch activates its Auto-IP configuration feature.

2 The Auto-IP feature starts with an IP address of 169.254.100.100. It uses the Address Resolution Protocol (ARP) to check to make sure this address is not already in use on the network. If not, it will allocate this default address to the Switch.

If this IP address is already in use, Auto-IP will check once every second for three seconds for an IP address on the 169.254.x.y subnet (where x = 1-254 and y = 0-255) (Auto-IP only uses addresses in the range 169.254.1.0 through to 169.254.254.255 as valid addresses.) Auto-IP uses the MAC address of the Switch as its starting point to produce an IP address for the unit. Once Auto-IP has ensured that an IP address is not already in use on the network, it assigns it to the Switch with a subnet mask of 255.255.0.0 and a default gateway of 0.0.0.0.

3 If DHCP or Auto-IP is not available, and the BOOTP server is configured, the IP address is assigned from that server.

4 While an Auto-IP assigned address is in use:

- The Auto-IP client continues to check every 30 seconds (using ARP) to ensure that any other Auto-IP hosts have not mistakenly configured themselves using the same Auto-IP address.
- DHCP and BOOTP requests also continue in the background. The requests begin 3 minutes after the Auto-IP address is assigned. The requests proceed with DHCP requests for 1 minute; a 3 minute pause; DHCP requests for another minute; a 3 minute pause; BOOTP requests

for one minute; a 3 minute pause; then the process repeats until a DHCP or BOOTP server answers the requests.

Important Considerations

This section contains some important points to note when using the automatic IP configuration feature.

The dynamic nature of automatically configured IP information means that a Switch may change its IP address whilst in use.

Server Support

Your Switch has been tested to interoperate with DHCP and BOOTP servers that use the following operating systems:

- Microsoft Windows 2000 Server
- Microsoft Windows NT4 Server
- Sun Solaris v2.5.1

If you want DHCP or BOOTP to be the method for automatic configuration, make sure that your DHCP or BOOTP servers are operating normally before you power on your Switch.

Event Log Entries and Traps

An event log will be generated and an SNMP trap will be sent if any of the following changes occur in the IP configuration:

- IP address configuration is changed manually
- IP address changes from Auto-IP to DHCP IP configuration
- DHCP negotiates a change in the IP configuration

To receive SNMP traps you must configure an SNMP trap destination. You can do this via the **`system management snmp trap`** *CLI command.*

10 IP ROUTING

Routing is a method for distributing traffic throughout an IP network. It is used to join LANs at the network layer (Layer 3) of the Open Systems Interconnection (OSI) model. A router provides both filtering and bridging functions across the network.

This chapter explains routers, protocols, and how your Switch allows bridges and routers to interoperate. It covers the following topics:

For detailed information on setting up your Switch for management, see the Getting Started Guide that accompanies your Switch.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM supplied with your Switch or on the 3Com Web site.

What is Routing?

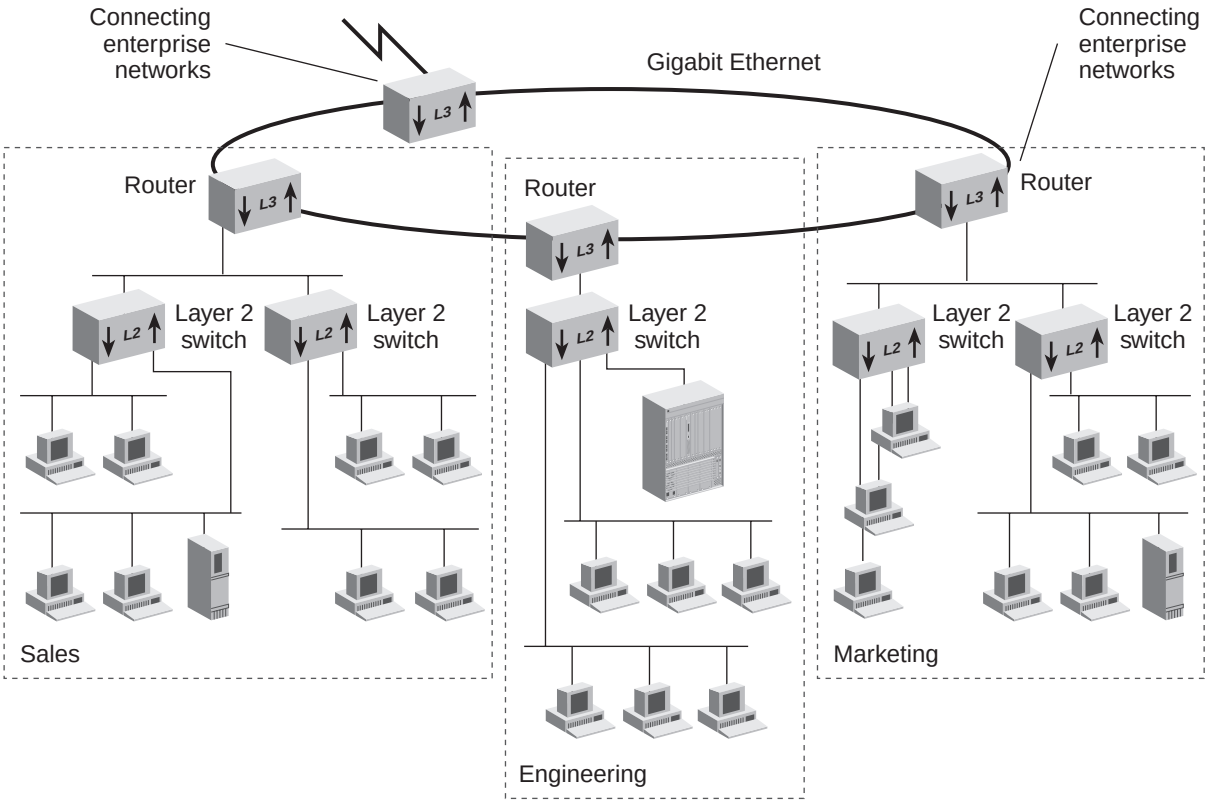
Routing distributes packets over potentially dissimilar networks. A router is the device that accomplishes this task. Your Switch, as a Layer 3 device, can act as a router. Routers typically:

- Connect enterprise networks.

- Connect subnetworks (or client/server networks) to the main enterprise network.

Figure 21 shows where routers are typically used in a network. Routing connects subnetworks to the enterprise network, providing connectivity between devices within a workgroup, department, or building.

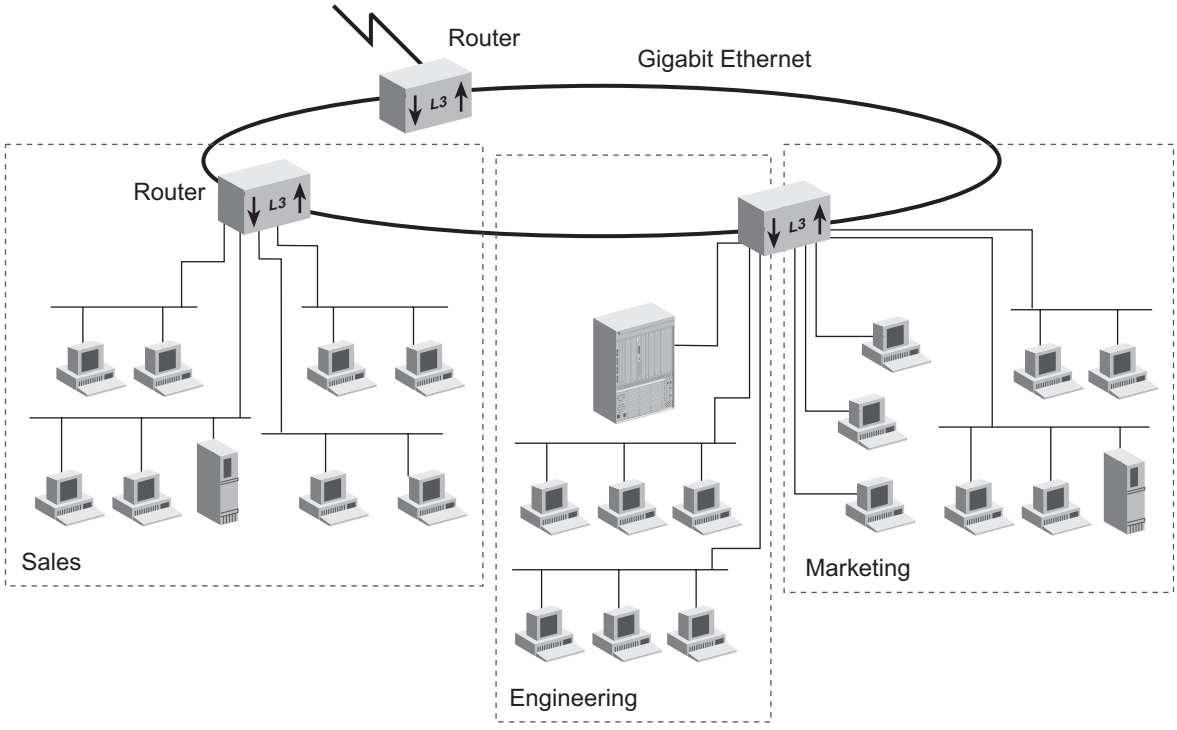
Figure 21 Typical Routing Architecture



Routing in a Subnetworked Environment

Your Switch allows you to both perform routing and switching within your network. You can streamline your network architecture by routing between subnetworks and switching within subnetworks. See Figure 22 for an example configuration.

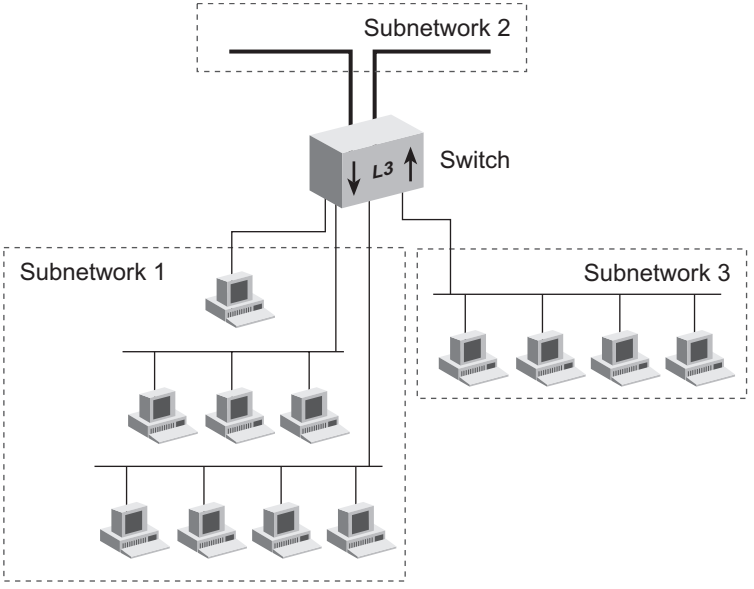
Figure 22 Subnetwork Routing Architecture



Integrating Bridging and Routing

Your Switch integrates bridging and routing. You can assign multiple ports to each subnetwork. See Figure 23 for an example configuration.

Figure 23 Multiple Ethernet Ports Per Subnetwork



Bridging switches traffic between ports that are assigned to the same subnetwork. Traffic traveling to different subnetworks is routed using one of the supported routing protocols.

Bridging and Routing Models

Your Switch implements routing differently from the way bridges and routers usually coexist.

Traditionally, network systems first try to route packets that belong to recognized protocols; all other packets are bridged.

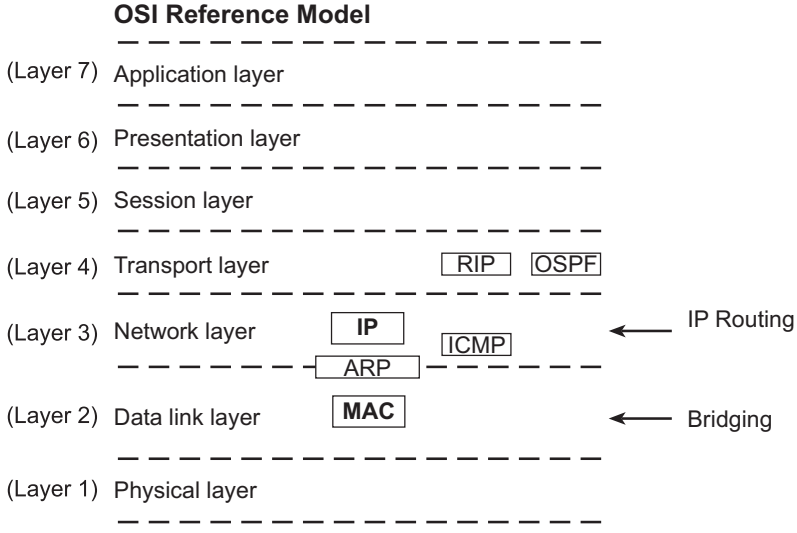
Your Switch first tries to determine if the packet is to be routed or switched. If the destination MAC address matches the MAC address of a router port on this Switch and the packet is not a protocol request to the Switch itself, then the packet is routed. However, if the destination MAC address is not the MAC address for a port on this Switch, the packet is further examined to determine if it can be switched according to the IEEE Std 802.1D, 1998 Edition.

Route calculations are triggered during initialization and when changes are made in the network configuration. At that point the router determines the appropriate route and forwards to the switch information describing the path that is to be taken. Once the route has been determined, all packets in the current flow are simply switched or forwarded across the chosen path. This takes advantage of the high throughput and low latency characteristics of switching by enabling the traffic to bypass the routing software once path calculation has been performed.

What is IP Routing?

An IP router, unlike a bridge, operates at the network layer of the OSI Reference Model. The network layer is also referred to as Layer 3. An IP router routes packets by examining the network layer address (IP address). Bridges use data link layer MAC addresses (at Layer 2) to perform forwarding. See Figure 24.

Figure 24 OSI Reference Model and IP Routing



When an IP router sends a packet, it does not know the complete path to a destination — only the next hop (the next device on the path to the destination). Each hop involves three steps:

1. The IP routing algorithm computes the *next hop* IP address and the next router interface, using routing table entries.
2. The Address Resolution Protocol (ARP) translates the next hop IP address into a physical MAC address.
3. The router sends the packet over the network across the next hop.

Benefits of IP Routing

IP routing provides the following features and benefits:

- **Economy** — Because you can connect several segments to the same subnetwork with routing, you can increase the level of segmentation in your network without creating new subnetworks or assigning new network addresses. Instead, you can use additional Ethernet ports to expand existing subnetworks.
- **Optimal routing** — IP routing can be the most powerful tool in a complex network setup for sending devices to find the best route to receiving devices. (The best route here is defined as the shortest and fastest route.)
- **Resiliency** — If a router in the network goes down, the other routers update their routing tables to compensate for this occurrence; in a typical case, there is no need for you to manually intervene.

IP Routing Concepts

IP routers use the following elements to transmit packets:

- Router Interfaces
- Routing Tables
- VLAN-based Routing
- Multiple IP Interfaces per VLAN

Router Interfaces

A router interface connects the router to a subnetwork. On your Switch, more than one port can connect to the same subnetwork.

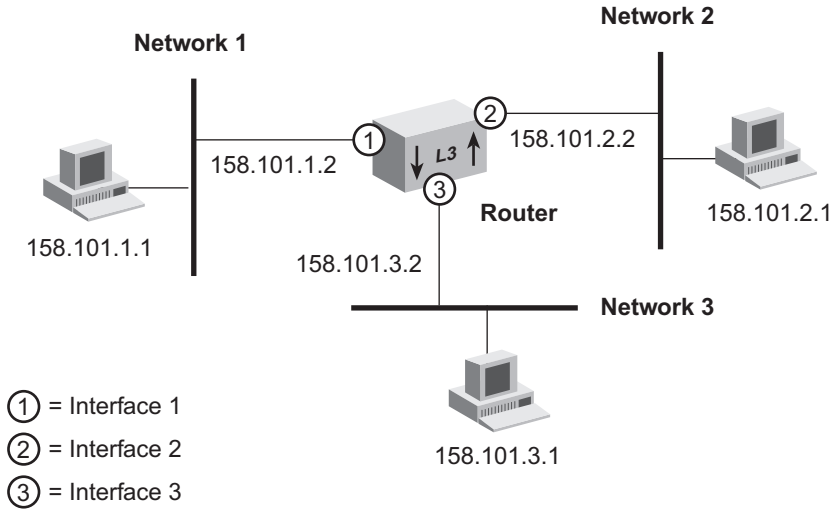
Each router interface has an IP address and a subnet mask. This router interface address defines both the number of the network to which the

router interface is attached and its host number on that network. A router interface IP address serves three functions:

- Sends IP packets to or from the router.
- Defines the network and subnetwork numbers of the segment that is connected to that interface.
- Provides access to the Switch using TCP/IP or to manage the Switch using the Simple Network Management Protocol (SNMP)

Figure 25 shows an example of a router interface configuration.

Figure 25 Routing Interfaces



Routing Tables

With a routing table, a router or host determines how to send a packet toward its ultimate destination. The routing table contains an entry for every learned and locally defined network. The size of the routing table is dynamic and can hold at most 2000 entries.

A router or host uses the routing table when the destination IP address of the packet is not on a network or subnetwork to which it is directly connected. The routing table provides the IP address of a router that can forward the packet toward its destination.

The routing table consists of the following elements:

- **Destination IP address** — The destination network, subnetwork, or host.
- **Subnet mask** — The subnet mask for the destination network.
- **Metric** — A measure of the distance to the destination. In the Routing Information Protocol (RIP), the metric is the number of hops through routers.
- **Gateway** — The IP address of the router interface through which the packet travels on its next hop.
- **Status** — Information that the routing protocol has about the route, such as how the route was put into the routing table.

Routing table data is updated statically or dynamically:

- **Statically** — You manually enter static routes in the routing table. You can define up to 100 (maximum) static routes. Static routes are useful in environments where no routing protocol is used or where you want to override some of the routes that are generated with a routing protocol. Because static routes do not automatically change in response to network topology changes, manually configure only a small number of reasonably stable routes. Static routes do not time out, but they can be learned.
- **Dynamically** — Routers use a protocol such as RIP or OSPF to automatically exchange routing data and to configure their routing tables dynamically. Routes are recalculated at regular intervals. This process helps you to keep up with network changes and allows the Switch to reconfigure routes quickly and reliably. Interior Gateway Protocols (IGPs), which operate within networks, provide this automated method.

Default Route

In addition to the routes to specific destinations, a routing table can contain a *default route*. The router uses the default route to forward packets that do not match any other routing table entry.

A default route is often used in place of static routes to numerous destinations that all have the same gateway IP address and interface number. The default route can be configured statically, or it can be learned dynamically.

A drawback to implementing a default static route is that it is a single point of failure on the network.

VLAN-based Routing

VLAN-based routing is used to control how a bridge and a router interact within the same Switch. The Switch uses a routing over bridging scheme, first trying to determine if the packet will be switched or routed. The Switch does this by examining the destination MAC address:

- If the destination MAC address is the internal router port on this Switch, the packet is routed (Layer 3).
- If the destination MAC address is not one of the router interfaces MAC addresses on this Switch, then the packet will be switched and is forwarded according to the IEEE 802.1D protocol.

This model allows the Switch to route the packet first, and then if the packet cannot be routed, give the packet to Layer 2 to be bridged by the VLAN. This scheme gives you the flexibility to define router interfaces on top of several bridge ports.

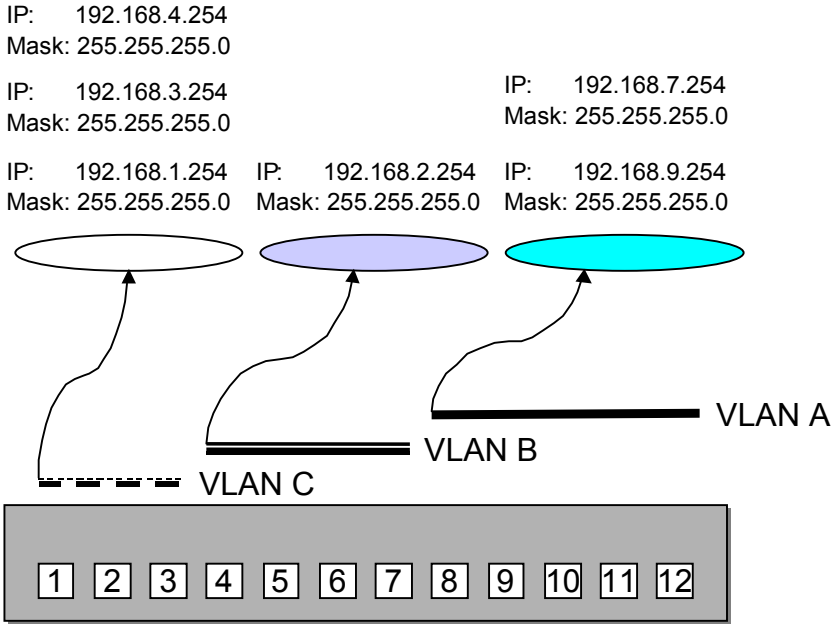
The "routing over bridging" scheme requires a VLAN-based IP Interface. To create this kind of interface you must first configure a VLAN and then create a router interface over that VLAN.

See Chapter 8 for more information on VLANs.

Multiple IP Interfaces per VLAN

You can overlap IP interfaces without configuring a separate VLAN for each subnet. Multiple IP interfaces can share the same VLAN, allowing multiple subnets to be routed on the same 802.1Q VLAN. You can define up to 64 IP interfaces on the Switch, that is, IP routing interfaces for static VLANs. See Figure 26.

Figure 26 Multiple IP Interfaces per VLAN



Implementing IP Routing

To route network traffic using IP, you must perform these tasks in the following order:

1. Configure Manual Aggregated Links (Optional)
2. Configure IP VLANs
3. Configure Automatic (LACP) Aggregated Links (Optional)
4. Establish IP Interfaces

If you are using XRN Technology on your network, 3Com recommends that you have LACP enabled on the Switches in the Distributed Fabric and that you omit step 1 and carry out steps 2, 3, and 4.

Configuring Manual Aggregated Links (Optional)

Aggregated Links work at Layer 2 and allow you to combine multiple Fast Ethernet or Gigabit Ethernet ports into a single high-speed link between two Switches.

If you intend to use manual aggregated links on an IP device, configure your aggregated links *before* you set up VLANs and IP interfaces. In this case, you must specify the index number of the aggregation. For example, if ports 5 through 8 are associated with an aggregated link, specifying "Aggregated Link 1" defines the VLAN to include all of the physical ports in the aggregation (ports 5 through 8).

Configuring IP VLANs

If you want to use IP routing, you must first configure the VLAN to use IP. You can create network-based VLANs that are grouped according to the IP network address and mask.

See Chapter 8 for more information on VLANs.

Configuring Automatic Aggregated Links (LACP)

If you intend to use automatic aggregated links (LACP) on an IP device, you must configure your VLANs *before* you connect your ports and set up IP interfaces.

Establishing IP Interfaces

To establish an IP interface:

1. Determine your interface parameters.
2. Define the IP interfaces.

Interface Parameters

Each IP routing interface has these standard characteristics:

- **IP address** — An address from the range of addresses that the Internet Engineering Task Force (IETF) assigns to your organization. This address is specific to your network and Switch. Refer to Appendix C for details on IP Addressing.
- **Subnet mask** — The 32-bit number that uses the same format and representation as an IP address. The subnet mask determines which bits in the IP address are interpreted as the network number/subnetwork number and the host number. Each IP address bit that corresponds to a 1 in the subnet mask is in the network/subnetwork part of the address. Each IP address bit that corresponds to a 0 is in the host part of the IP address.
- **State** — The status of the IP interface. It indicates whether the interface is available for communications (`Up`) or unavailable (`Down`). This is not a user configurable parameter.
- **VLAN interface index** — The number of the VLAN that is associated with the IP interface. When the Switch prompts you for this option, the menu identifies the available VLAN indexes.

Important Consideration

Consider the following issue before you establish an IP interface:

- Before you assign IP addresses, map out the entire network and subnetwork IP addressing scheme. Plan for future expansion of address numbers as well.

Defining an IP Interface

After you determine the VLAN index, IP address, and subnet mask for each IP interface, you can define each interface. Use the Command Line Interface or the Web interface to define an IP interface.

Remember that you must define a VLAN before you define the IP (routing) interface. See Chapter 8 for more information on VLANs..

To define your IP interface, you should understand the following IP features:

- ARP Proxy
- ICMP Router Discovery

- Routing Information Protocol (RIP)
- User Datagram Protocol (UDP) Helper

These features are discussed later in this chapter.

You can use the Routing Information Protocol (RIP) protocol to take advantage of routing capabilities. RIP is discussed in this chapter.

Administering IP Routing

Keep these points in mind while you administer the IP network:

- Flush the ARP cache regularly if you set the age time to 0.
- Set up a default route.

  The Switch uses the default route to forward packets that do not match any other routing table entry. You may want to use the default route in place of routes to numerous destinations that all have the same gateway IP address. If you do not use a default route, ICMP is more likely to return an `ICMP Network Unreachable` error.
- Before you can define static routes, you must define at least one IP interface. See "Defining an IP Interface" on page 102 for more information. Remember the following guidelines:
  - Static routes remain in the routing table until you remove them or the corresponding interface.
  - Static routes are removed during a system initialize.
  - Static routes take precedence over dynamically learned routes to the same destination.
  - Static routes are included in periodic RIP updates sent by your Layer 3 Switch.

IP Routing Protocols

IP protocols are a set of uniquely defined interactions that allow data communications to occur. Protocols are the rules to which networks must adhere in order to successfully operate. Protocols that are discussed in this section include:

- Address Resolution Protocol (ARP)
- Internet Control Message Protocol (ICMP)
- Routing Information Protocol (RIP)
- User Datagram Protocol (UDP) Helper

Address Resolution Protocol (ARP)

ARP is a low-level protocol that locates the MAC address that corresponds to a given IP address. This protocol allows a host or router to use IP addresses to make routing decisions while it uses MAC addresses to forward packets from one hop to the next.

You do not need to implement ARP — the Switch has ARP capability built in, but you can change and display the contents of the ARP cache.

When the host or router knows the IP address of the *next* hop towards the packet destination, the host or router translates that IP address into a MAC address before sending the packet. To perform this translation, the host or router first searches its *ARP cache*, which is a table of IP addresses with their corresponding MAC addresses. Each device that participates in IP routing maintains an ARP cache. See Figure 27.

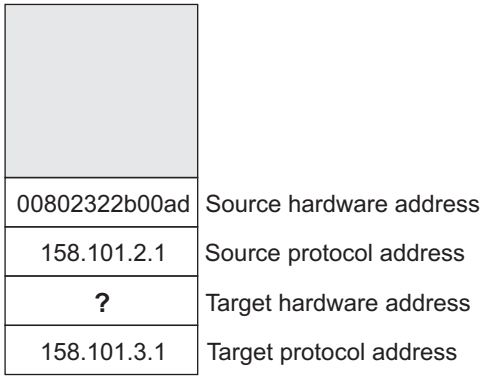
Figure 27 Example of an ARP Cache

ARP cache

IP address	MAC address
158.101.1.1	00308e3d0042
158.101.2.1	0080232b00ab

If the IP address does not have a corresponding MAC address, the host or router broadcasts an *ARP request* packet to all the devices on the network. The ARP request contains information about the target and source addresses for the protocol (IP addresses). See Figure 28.

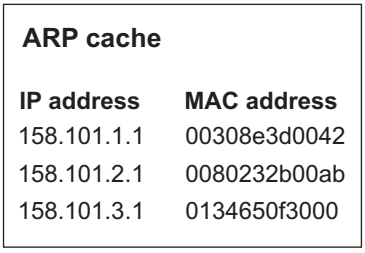
Figure 28 Example of an ARP Request Packet



When devices on the network receive this packet, they examine it. If their address is not the target protocol address, they discard the packet. When a device receives the packet and confirms that its IP address matches the target protocol address, the receiving device places its MAC address in the target hardware address field and exchanges both source and target fields. This packet is then sent back to the source device.

When the originating host or router receives this *ARP reply*, it places the new MAC address in its ARP cache next to the corresponding IP address. See Figure 29.

Figure 29 Example of ARP Cache Updated with ARP Reply



After the MAC address is known, the host or router can send the packet directly to the next hop.

ARP Proxy

ARP proxy allows a host that has no routing ability to determine the MAC address of a host on another network or subnet.

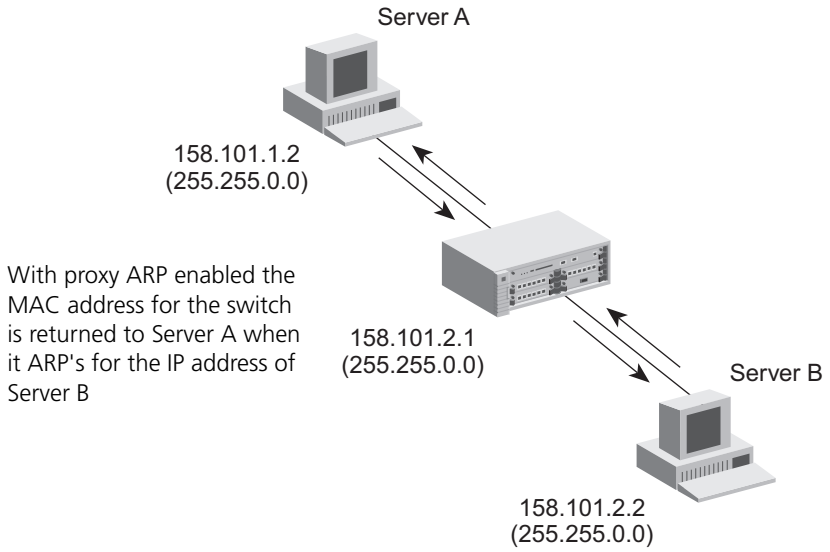
When ARP proxy is enabled and a workstation sends an ARP request for a remote network, the Switch determines if it has the best route and then answers the ARP request by sending its own MAC address to the workstation. The workstation then sends the frames for the remote destination to the Switch, which uses its own routing table to reach the destination on the other network.

Example

In the following example, Server A cannot use the router as a gateway to Server B because Server A has its subnet mask set to broadcast (using ARP) its IP network address as 158.101.0.0, while the IP network address of the router is 158.101.1.0.

However, if the router answers the request of Server A with its own MAC address — thus, all traffic sent to Server B from Server A is addressed to the corresponding IP interface on the router and forwarded appropriately.

Figure 30 ARP Proxy



Internet Control Message Protocol (ICMP)

Because a router knows only about the next network hop, it is not aware of problems that may be closer to the destination. Destinations may be unreachable if:

- Hardware is temporarily out of service.
- You specified a nonexistent destination address.
- The routers do not have a route to the destination network.

To help routers and hosts discover problems in packet transmission, a mechanism called Internet Control Message Protocol (ICMP) reports errors back to the source when routing problems occur. With ICMP, you can determine whether a delivery failure resulted from a local or a remote problem.

ICMP performs these tasks:

- **Determines which router to use as the default gateway (*ICMP Router Discovery*)** — ICMP Router Discovery is useful if you have

multiple gateways that connect a particular subnet to outside networks.

ICMP Router Discovery

ICMP Router Discovery enables hosts that are attached to multicast or broadcast networks to discover the IP addresses of their neighboring routers and determine which router to use for a default gateway.

ICMP Router Discovery is permanently enabled on the Switch.

Important Considerations

Keep the following points in mind with ICMP Router Discovery:

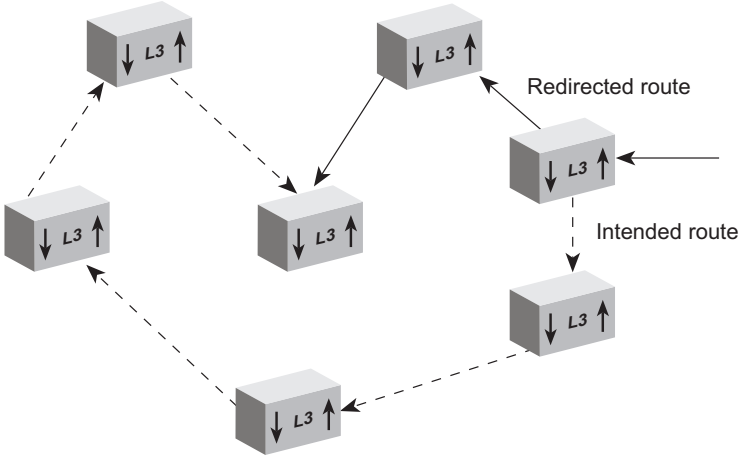
- ICMP Router Discovery is useful on large networks, or when the network topology has undergone a recent change.
- If you are on a small network that is relatively stable, consider using a static route to the gateway instead of ICMP Router Discovery to reduce network traffic.

See the documentation for your workstation to determine whether you can configure your workstation to use this protocol.

See RFC 1256 for detailed information about ICMP Router Discovery.

Figure 31 shows how ICMP can dynamically determine a router to act as the default gateway.

Figure 31 ICMP Router Discovery



Routing Information Protocol (RIP)

RIP is the protocol that implements routing. RIP does this by using Distance Vector Algorithms (DVAs) to calculate the route with the fewest number of hops to the destination of a route request. Each device keeps its own set of routes in its routing table. RIP is an Interior Gateway Protocol (IGP) for TCP/IP networks.

RIP operates using both active and passive devices.

- *Active* devices, usually routers, broadcast RIP messages to all devices in a network or subnetwork and update their internal routing tables when they receive a RIP message.
- *Passive* devices, usually hosts, listen for RIP messages and update their internal routing tables, but do not send RIP messages.

An active router sends a broadcast RIP message every 30 seconds. This message contains the IP address and a metric (distance) from the router to each destination in the routing table. In RIP, each router through which a packet must travel to reach a destination counts as one network *hop*.

Basic RIP Parameters

There are several parameters to consider when you set up RIP for your network. When you configure an IP interface, the Switch already has the RIP parameters set to the defaults listed in Table 14.

Table 14 RIP Parameters

RIP Parameter	Default Value
Router Mode*	disabled
Cost†	1
Update Time*	30 seconds
Send Mode	RIPv2Compatible
Receive Mode	RIPv2OrRIPv2
Poison Reverse	disabled
Advertisement Address	limited broadcast address (224.0.0.9)

* These RIP parameters apply to the entire Switch. All other parameters are defined per interface.

† The Cost value cannot be altered, it is fixed at 1.

Router Mode

The available settings for router mode are as follows:

- **Disabled** — The Switch ignores all incoming RIP packets and does not generate any RIP packets of its own.
- **Enabled** — The Switch broadcasts RIP updates and processes incoming RIP packets.

Update Time

This Switch sends a RIP message every 30 seconds (by default) with both the IP address and a *metric* (the distance to the destination from that router) for each destination. You can modify the update time if needed to adjust performance.

Send and Receive Modes

The following RIP send and receive modes are supported by the Switch:

Table 15 Send and Receive Modes

Send Mode	Receive Mode
RIPv1	RIPv1
RIPv1Compatible	RIPv2
RIPv2	RIPv1OrRIPv2
doNotSend	doNotReceive

- RIPv1 – Route information is broadcast periodically to other routers on the network using the advertisement list for RIP-1 updates.
- RIPv2 – Route information is multicast periodically to other routers on the network using the multicast address of 224.0.0.9. This method reduces the load on hosts that are not configured to listen to RIP-2 messages.
- RIPv1 Compatible – Route information is broadcast to other routers on the network using the advertisement list for RIP-2 updates.
- RIPv1OrRIPv2 – Both RIP-1 and RIP-2 route information can be received by the Switch.
- doNotSend – The Switch processes (or passively learns) all incoming RIP packets, but does not transmit RIP updates.
- doNotReceive – The Switch broadcasts (or advertises) RIP updates, but does not process incoming RIP packets.

The doNotSend and doNotReceive modes are also referred to as one-way learn and advertise modes.

Poison Reverse

Poison Reverse is a RIP feature that you use specifically with a scheme called *Split Horizon*. The Switch disables Poison Reverse by default.

Split Horizon avoids the problems that reverse-route updates can cause. Reverse-route updates are sent to a neighboring router and include the routes that are learned from that router. Split Horizon omits the routes that are learned from one neighbor in the updates that are sent to that neighbor (the reverse routes).

Poison Reverse is essentially another layer of protection against advertising reverse routes.

- When you enable Poison Reverse, the Switch advertises reverse routes in updates, but it sets the metrics to 16 (infinity). Setting the metric to infinity breaks the loop immediately when two routers have routes that point to each other.
- When you disable (default mode) Poison Reverse, such reverse routes are not advertised.

You can disable Poison Reverse because it augments what Split Horizon already does, and it puts additional information that you may not need into RIP updates.

Advertisement Address

The Switch uses the advertisement address to advertise routes to other stations on the same network. Each interface that you define uses a directed broadcast address as the advertisement address. The Switch uses this address for sending updates.

RIP-1 Versus RIP-2

Like RIP-1, RIP-2 allows the Switch to dynamically configure its own routing table. RIP-2 is much more flexible and efficient than RIP-1, however, because RIP-2 advertises using the multicast method, which can advertise to a subset of the network (RIP-1 uses the broadcast method, which advertises to the whole network). RIP-2 can do this because it includes a subnet mask in its header.

If your Switch receives a RIP-2 packet, your Switch puts the route into the routing table with the subnet mask that was advertised.

Important Considerations

Note the following considerations when you implement RIP on your Switch:

- Use RIP-2 rather than RIP-1 if possible, because RIP-2 uses subnet masking and the next hop field. Subnet mask advertising allows you to use VLSM (Variable Length Subnet Mask).
- Where possible, set RIP as follows:
  - **Send Mode** — `RIPv2`
  - **Receive Mode** — `RIPv1OrRIPv2`

  In this way, the Switch keeps track of the RIP-1 and RIP-2 address routes in its routing table and forwards the routes as well.
- When using Spanning Tree (STP), Rapid Spanning Tree (RSTP) and Routing Information Protocol (RIP) all Switches must communicate with each other on the same VLAN.

User Datagram Protocol (UDP) Helper

User Datagram Protocol (UDP) Helper allows TCP/IP applications to forward broadcast packets from one part of the network to another. The most common uses of UDP are:

- **Bootstrap Protocol (BOOTP)**

  BOOTP allows you to boot a host through the router using a logical port. This can be done even when the host is on another part of the network. UDP packets that rely on the BOOTP relay agent are modified and then forwarded through the router.
- **Dynamic Host Configuration Protocol**

  A host can retrieve its own configuration information including IP address, from a DHCP server through the IP network. DHCP makes it easier to administer the IP network. With DHCP you can dynamically configure a host with new information.

Implementing UDP Helper

Your Switch implements a generic UDP Helper agent that applies to any port. You have to set the following UDP Helper parameters:

- **UDP Port Number** A logical address, not a port (interface) on your device. BOOTP (including DHCP) uses UDP port 67.
- **IP forwarding address** The IP address to which the packets are forwarded. You can have up to 32 combinations of port numbers and IP forwarding addresses. You can also have up to 4 IP address entries for the same ports.

You need to have a thorough understanding of your network configuration to use UDP Helper. Review the network topology before you implement UDP Helper.

Important Considerations

Consider the following points when you use UDP Helper:

- Overlapped IP interfaces are multiple logical interfaces that are defined for a single VLAN. When forwarding BOOTP/DHCP packets, UDP Helper includes the source interface in the forwarded packet. In the case of overlapped IP interfaces, UDP Helper will rotate through the multiple interfaces in a round-robin fashion when inserting a source interface into the forwarded packet. This allows the DHCP server to distribute addresses over all of the interfaces. If the originating host requests an address on a specific interface, than UDP Helper will use that interface, overriding the round-robin process.
- The BOOTP hop count (how many steps the Switch uses to forward a packet) is fixed at 16.
- You can always add or remove a port number or IP forwarding address defined for UDP Helper.

Advanced IP Routing Options

Your Switch has several features which further extend the networking capabilities of the device. Refer to Appendix D for more information on the following:

- Variable Length Subnet Masks (VLSMs)
- Supernetting

Access Control Lists

Access Control Lists are a set of instructions that can be applied to filter traffic on VLANs. They can be used to limit access to certain segments of the network and therefore, are useful for network security.

Access Control Lists can be used to:

- Prevent unnecessary network traffic.
- Restrict access to proprietary information within the network.

Access Control Lists are based on a series of rules. Rules are applied to VLANs and determine the path or access limitations for packets received on a VLAN. When a packet is received on a VLAN, it is compared to an access list for this VLAN. If a match is found; meaning the packet falls under the rule, it will be blocked or forwarded to the appropriate VLAN depending on the action.

Rules are established based on IP addressing. A packet matches an access list rule when it's destination IP address falls with the values of the rule. When a match is found, the path the packet takes is determined by the rule and is either forwarded (permitted) or dropped (denied).

There are a maximum of 100 access lists that can be applied under the current operating system. Access list rules can be applied and traffic is forwarded at wire speed using layer 3 destination IP addresses and VLANs.

How Access Control List Rules Work

When a packet is received it is compared against the VLAN access list. The access list rules are applied to a range of IP addresses and are defined by the destination IP address and a mask. If a match is found in the access list the appropriate action is taken. By default, if no access list has been defined for a VLAN, all IP traffic will be permitted. Denial is based on a pre-defined rule.

For example:

Packet destination IP address: 10.101.67.45

Rule destination address: 10.101.67.0

Rule destination mask: 255.255.255.0

Rule action: deny

As a result of the above rule, the packet matches the parameters of the rule and will be blocked.

A destination mask of 0.0.0.0 will match all packets.

11 USING WEBCACHE SUPPORT

This chapter outlines the Webcache support feature, explains the key benefits of using this feature, and gives examples of how and why you would use it in your network.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM that accompanies your Switch or on the 3Com Web site.

What is Webcache Support?

Webcache support is a feature that allows local storage (caching) of frequently accessed web pages on a Webcache attached to your network. This means your network users can access these locally stored web pages without going over a WAN connection.

The Webcache periodically checks live web pages to find out if the current cached pages are out-of-date and replaces them accordingly.

Supported Devices

The Webcache feature is available on the following Switches:

- SuperStack 3 Switch 4924 (3C17701)
- SuperStack 3 Switch 4950 (3C17706)
- 3Com Switch 4050 (3C17708)
- 3Com Switch 4060 (3C17709)

Benefits of Webcache Support

The primary benefit of the Webcache support feature is to increase the performance of the local network by redirecting HTTP (web) traffic to a local Webcache. An increase in network performance is achieved because:

- traffic on a WAN connection is reduced as the local cache, rather than remote web servers can serve requests from multiple users that are accessing the same web content.
- latency is reduced as the Webcache is able to deliver web content faster than the time required to retrieve information over a WAN connection.

Because the redirection decision is based upon the destination TCP port 80, the solution is transparent to end users and requires no manual configuration of web clients.

How Webcache Support Works

The Webcache feature enables the Switch to redirect web-based (HTTP) traffic to a Webcache. Redirecting HTTP traffic reduces traffic on the WAN connection because a local cache services the requests from multiple users rather than remote web servers. Latency is reduced because the cache is able to deliver web content faster than a WAN connection.

Configuration of a Webcache on the Switch is very flexible. It can be configured to any VLAN as long as there is an IP interface associated with the VLAN. HTTP traffic from any VLAN can be redirected to the webcache. To operate properly, all VLANs that access the Webcache must have their own IP interface defined. Alternatively, if you should choose not to have HTTP traffic directed to an IP address you can prevent it by using the IP Exclusion option. Refer to "IP Exclusions" on page 120 for more information.

Cache Health Checks

The cache health check is a feature that ensures web traffic is not redirected to a cache that is not currently operating. The health check works as follows:

1. The health check requests a factory-defined URL from the Webcache every eleven seconds and expects to receive a reply to confirm that the cache is operating normally.
2. If a reply is not received from the Webcache, the Switch will start polling the Webcache at three second intervals.
3. If the Webcache fails three health check attempts, the Webcache is deemed to have failed and the Webcache support feature on the Switch is disabled (that is, it no longer redirects HTTP traffic). From this point on all HTTP traffic will go directly to the WAN.

4 However, the Webcache support feature, although no longer redirecting traffic, continues to perform health checks on the Webcache at three second intervals to determine if the Webcache is operating. If a health check is successful, redirection of HTTP traffic will start again.

Webcache Support Examples

The following examples explain how Layer 4 redirection operates over a Webcache in two different VLAN configurations.

Bridging Over A Single VLAN Example

When the endstation, Webcache and Router are all in the same VLAN, Layer 4 Redirection can take place by bridging at Layer 2. In the example shown in Figure 32:

1 An endstation sends a request packet to the Switch. The request packet contains the Web server's IP address as its destination. The TCP destination port is 80 (the default for an HTTP request).

2 The Switch detects that the incoming packet should be redirected. It sends the packet to the Webcache.

3 The Webcache either:

- Sends the cached page to the endstation, if the page is already cached. Routing is not required because the endstation and the Webcache are both in the same VLAN

or

- If the Webcache does not have the requested page already cached:

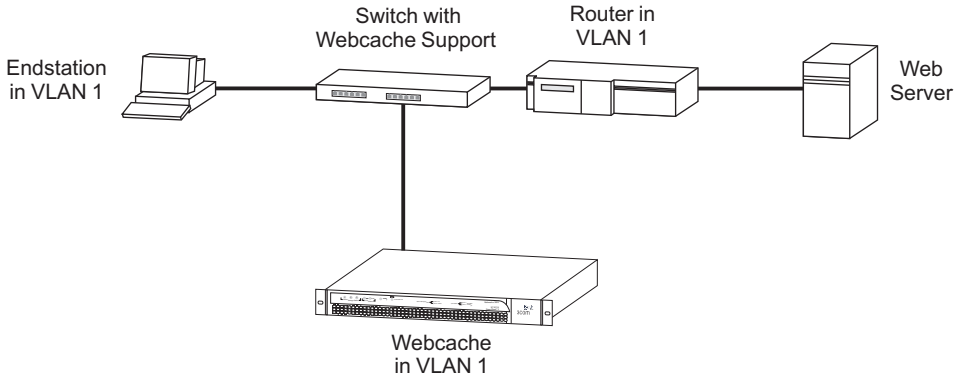
a The Webcache forwards the HTTP request to the Web server, using the destination IP address supplied by the endstation's original request. This request uses the MAC address and IP address of the Webcache as its source.

b The Web server sends the requested page back to the Webcache. The destination TCP port of the requested page is a random port number other than 80. The Switch treats this packet in the same way as any other packet, and sends it to the Webcache over a bridge. Routing is not required because the Router and Webcache are both in the same VLAN.

c The Webcache sends the page to the requesting endstation over a bridge.

In this example configuration, the traffic being received and transmitted by the Webcache is bridged. However, an IP interface must always be defined on a VLAN containing a Webcache. This is to ensure that the Webcache is configured correctly, and to ensure that the periodic health checks performed by Layer 4 redirection are successful.

Figure 32 Webcache and Layer 4 Traffic Redirection on the same VLAN



Routing Over Multiple VLANs Example

When the endstation, Webcache and Router are in two or more different VLANs, Layer 4 redirection requires that IP interfaces are defined for each VLAN. Other than this, the operations are similar to those described in the "Bridging Over A Single VLAN Example".

In the example shown in Figure 33:

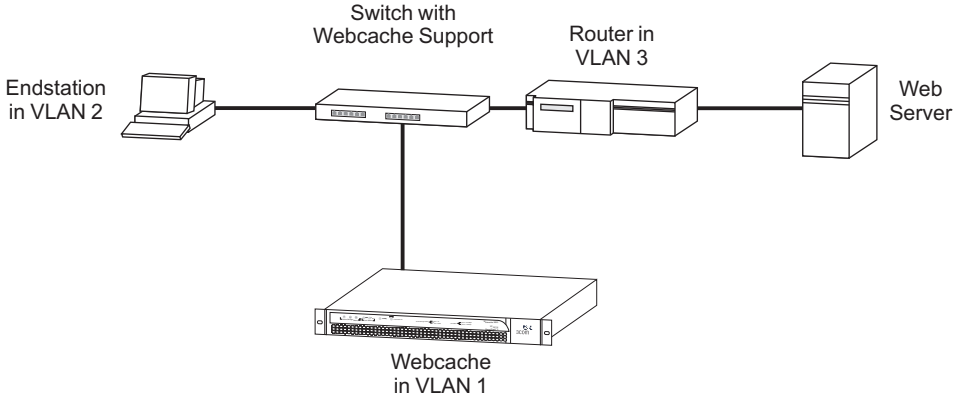
1. An endstation sends a request packet to the Switch. The request packet contains the Web server's IP address as its destination. The TCP destination port is 80 (the default for an HTTP request).
2. The Switch detects that the incoming packet should be redirected. It sends the packet to the Webcache. The packet must be routed because the endstation and the Webcache are on different VLANs.
3. The Webcache either:
   - Sends the cached page to the endstation, if it has the page already cached.

   or:

- If the Webcache does not have the requested page already cached, it forwards the request to a Web server. The Web page is then returned to the Webcache. When the Webcache receives the requested page, it forwards it to the requesting endstation.

In this example configuration, the traffic being received and transmitted by the Webcache is routed.

Figure 33 Webcache and Layer 4 Traffic Redirection on different VLANs



Important Considerations

This section contains some important considerations when using Webcache support on the Switch.

- The Switch supports the SuperStack 3 Webcache 1000/3000.
- The Webcache must be connected directly to the Switch — there must be no intervening Switches or Hubs.
- The Switch can only support one Webcache for a single unit or a Distributed Fabric.
- The traffic between any two pairs of IP addresses must always be redirected through the same Webcache.
- The port to which the Webcache is connected cannot be a member of an aggregated link.
- IP packets with IP Options set will not be redirected.

IP Exclusions

The IP Exclusion feature of the Switch simplifies Webcache administration.

IP Exclusions are used to identify HTTP traffic which you do not want to redirect. An IP Exclusion can be a single IP Address or an entire IP network depending on the IP Address mask that you configure.

The Switch looks at the Destination IP Address of the HTTP request, and if it matches a configured IP Exclusion, the traffic will not be directed to the Webcache, but will be switched normally.

IP Exclusions would be best or most commonly used for Intranet servers.

12 MAKING YOUR NETWORK SECURE

This chapter outlines the Switch Management Login feature, explains the key benefits of using this feature, and gives examples of how and why you would use it in your network.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM that accompanies your Switch.

What is Switch Management Login?

If you intend to manage the Switch using the Web interface or the command line interface, you need to log in with a valid user name and password.

For further information on managing the Switch, see the "Setting Up For Management" chapter in the Switch Getting Started Guide.

The user name and password information can be stored in either:

- **a RADIUS server** (recommended)

  If you enable RADIUS as the authentication mode of Switch Management Login, the user name and password information is stored in a database on a RADIUS server in your network. Subsequent log in attempts to the Switch are remotely authenticated by the RADIUS server.

  or

- **the local Switch database** (default)

  If you enable Local as the authentication mode of Switch Management Login, the user name and password information is stored in the local database on the Switch. Subsequent login attempts to the Switch are authenticated by the local database.

Benefits of RADIUS Authentication

Day-to-day network maintenance can become a substantial overhead. For example, regularly changing the administrative password on a manageable network device is a commonplace security measure. If the local Switch database is enabled, the network administrator must have local access to each Switch to securely change user name and password information. This can be time consuming, tedious and often results in bad configurations and lapses in security.

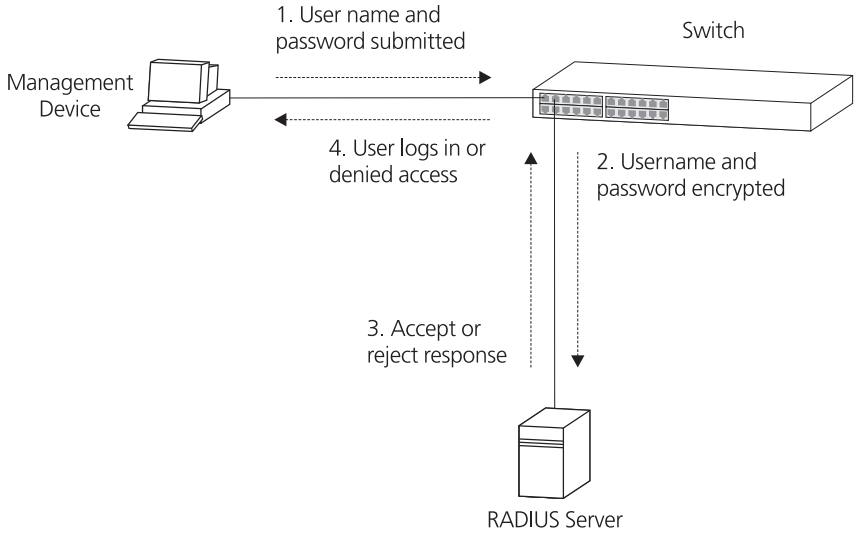
RADIUS authentication provides centralized, secure access and removes the need to physically visit each network device. Changes to user names and passwords require only a single action on the RADIUS database and are reflected immediately.

Your Switch is fully compliant with the industry standard RADIUS protocol. For further information about RADIUS, see "What is RADIUS?" on page 125.

How RADIUS Authentication Works

When RADIUS authentication of Switch Management Login is enabled, the Switch obtains the user's name and password and securely sends the information to the RADIUS server. The information is authenticated by the server and a valid user is allowed to login to the Switch. An invalid user will receive a reject response and is not allowed to login to the Switch. This process is shown in Figure 34.

Figure 34 RADIUS Authentication Operation



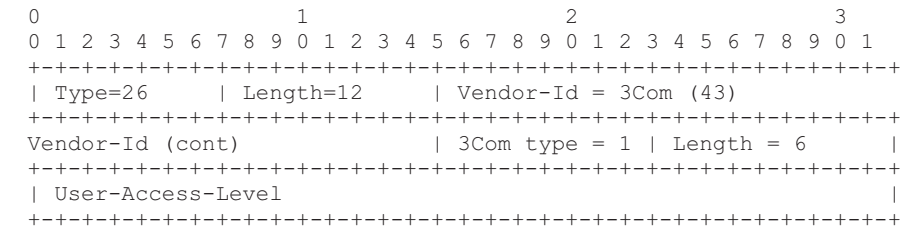
3Com Vendor Specific Attribute

The default user levels on the Switch (monitor, manager, admin) are supported by a 3Com Vendor Specific Attribute (VSA). The Vendor-ID for 3Com is 43. You must configure the RADIUS server to send this attribute in the Access-Accept message in order to specify the access level required for each user account. The configurable attribute values are:

- Monitor (1) — the user can view all manageable parameters, except special/security features, but cannot change any manageable parameters.
- Manager (2) — the user can access and change the operational parameters but not special/security features.
- Administrator (3) — the user can access and change all manageable parameters.

The attribute body consists of a 3Com Vendor type (1), Vendor data length (6) and the Vendor data (4 octet integer containing the access level value), as shown in Figure 35.

Figure 35 3Com Vendor Specific Attribute



For further information about configuring the Switch for Switch Management Login and RADIUS authentication, please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM that accompanies your Switch.

Important Considerations

This section contains some important considerations when using RADIUS authentication of Switch Management Login on the Switch.

- Before you enable RADIUS authentication you must ensure that:
  - The Switch is configured with a static IP address.
  - RADIUS has been configured on the Switch.
  - The RADIUS server in your network is operational.
  - The RADIUS server has been configured with the 3Com Authentication Vendor Specific Attribute (VSA). The Vendor-ID for 3Com is 43.
- If the Switch is unable to contact the RADIUS server, the command line interface automatically reverts to using the local Switch database for user authentication. This allows a user with "admin" access to login to the Switch via the console port and continue to manage it. The Web interface and Telnet do not revert to the local database, and the user will not be able to log in to the Switch via the Web interface or Telnet.
- The user names and passwords stored in the local Switch database may not be the same as those stored on the RADIUS server. When a user account is created on a RADIUS server, an equivalent account is not automatically created in the local Switch database, and vice versa.
- In routed environments, you must configure on the RADIUS server each IP interface on the Switch that may be used to connect to the RADIUS server.

What is RADIUS?

Remote Authentication Dial-In User Service (RADIUS) is an industry standard protocol for carrying authentication, authorization and configuration information between a network device and a shared authentication server. Transactions between each network device and the server are authenticated by the use of a shared secret. Additional security is provided by encryption of passwords to prevent interception by a network snooper.

RADIUS is defined in the RFC 2865 , "Remote Authentication Dial-in User Service (RADIUS)".

Switch Management Login, used to control administrative access, utilizes the RADIUS protocol.

13 3COM XRN TECHNOLOGY

This chapter explains what 3Com XRN Technology is and how you can use it to benefit your network. It also explains how to implement XRN on your network.

This chapter contains the following sections:

The sections below provide supplementary information that are not essential reading, but may be of interest to advanced users.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM that accompanies your Switch or on the 3Com Web site.

What is XRN Technology?

XRN (eXpandable Resilient Network) is a 3Com LAN core technology built into the software and hardware of your Switch that allows you to interconnect two Switches to create a Distributed Fabric. The interconnection enables the two Switches to operate as a single unit. This Distributed Fabric can be managed using a single IP address, with all the switching and routing distributed across the devices that make up the Distributed Fabric.

XRN technology provides a highly resilient core around which you can build your network.

Supported Switches

The XRN techology is supported on the following Switches installed with Version 3.0 software or later:

- SuperStack 3 Switch 4900 (3C17700)
- SuperStack 3 Switch 4900 SX (3C17702)
- SuperStack 3 Switch 4924 (3C17701)
- SuperStack 3 Switch 4950 (3C17706)
- 3Com Switch 4050 (3C17708)
- 3Com Switch 4060 (3C17709)

XRN Terminology

This section contains a glossary of the common XRN terminology.

eXpandable Resilient Network (XRN)

XRN is a technology developed by 3Com that allows you to implement fault tolerant, high performance and scalable multilayer Gigabit backbones on your network.

XRN Distributed Fabric (Distributed Fabric)

XRN Distributed Fabric is the term used to describe two interconnected devices supporting XRN technology.

XRN Interconnect

XRN Interconnect is the interconnection between two Switches supporting XRN technology that form the Distributed Fabric.

Distributed Device Management (DDM)

DDM allows both Switches in the XRN Distributed Fabric to behave as a single managed entity, irrespective of the form factor or Switch deployed. For further information see page 129.

Distributed Link Aggregation (DLA)

DLA is the configuration of Aggregated Links across both interconnected devices in the Distributed Fabric. 3Com and non-3Com devices can connect to the XRN Distributed Fabric using DLA. For further information see page 131.

Distributed Resilient Routing (DRR)

DRR provides distributed routing and router resiliency across an XRN Distributed Fabric. For further information see page 130.

Benefits of XRN Technology

The benefits of XRN technology include:

- Increased resilience provided by:
  - Hardware and Software redundancy per unit or across the Distributed Fabric.
  - Distributed management across the Distributed Fabric.
  - Distributed Link Aggregation across the Distributed Fabric.
  - Distributed Resilient Routing across the Distributed Fabric.
- Increased network performance provided by:
  - Multilayer switching capacity in excess of 80 million packets per second.
  - Up to 48 Gigabit ports across two units in a Distributed Fabric.
  - Link Aggregation supported across the Distributed Fabric.
- Flexibility provided by:
  - Support across Switch 4900 Series and Switch 4050/4060. You can mix and match any of the supported Switches in an XRN Distributed Fabric.

XRN Technology Features

This section describes the key features of XRN technology.

Distributed Device Management (DDM)

DDM provides single IP address management across the interconnected Switches that form the Distributed Fabric. This allows the entire Distributed Fabric to be managed and configured as a single managed entity. In the event of failure in one of the Switches in the Distributed Fabric, management access to the remaining Switch is retained on the same IP address.

DDM allows you to manage the Distributed Fabric via the command line interface (CLI), Web interface, or SNMP.

DDM provides you with the ability to carry out the following:

- Single step Switch software upgrades across the Distributed Fabric (provided the Switches are of the same family).
- Distributed Fabric-wide configuration of all software features.
- Configuration of port-specific software features across the Distributed Fabric via a single management interface.

Distributed Resilient Routing (DRR)

DRR allows the Switches in the Distributed Fabric to act as a single logical router which provides router resiliency in the event of failure in one of the interconnected Switches. With DRR, both Switches in the Distributed Fabric are routing, which significantly increases the overall Layer 3 capacity of the core of the network.

DRR can intelligently distribute the routing load across both Switches in the Distributed Fabric, which maximizes routing performance and makes full use of bandwidth capacity.

Switches in the Distributed Fabric provide Layer 3 local forwarding for directly connected hosts and devices.

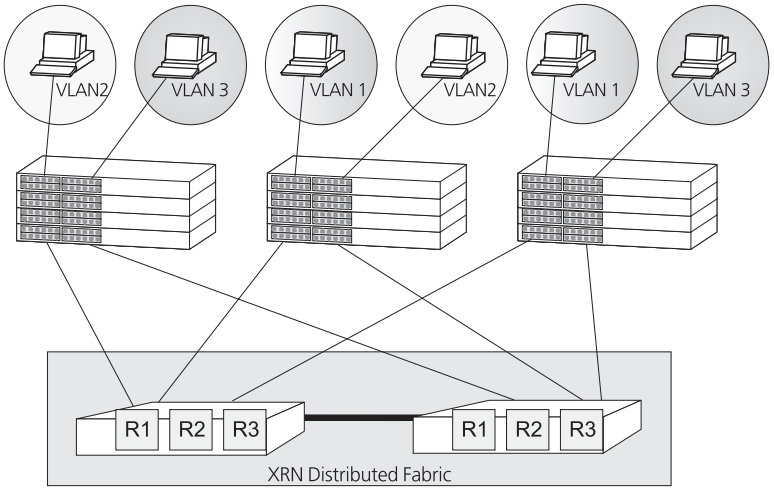
Both Switch units within the Distributed Fabric provide the same router interfaces and mirror each other's routing tables. This allows each unit to keep the routing local to the unit for locally connected hosts and devices.

In the example shown in Figure 36, there is a single logical router across the XRN Distributed Fabric with router interfaces (R1, R2, and R3) shared by both units.

If there is a loss of a unit in an XRN Distributed Fabric it does not affect routing provided you are using Distributed Link Aggregation or the devices in the wiring closet are multihomed.

DRR is an XRN-specific implementation that only operates on XRN technology within the Distributed Fabric. However it will interoperate with other routers outside of the XRN Distributed Fabric.

Figure 36 Network Example illustrating Distributed Resilient Routing



Distributed Link Aggregation (DLA)

DLA ensures that all member ports of an aggregated link distribute the traffic flow across the Distributed Fabric. This provides resilience and enhanced performance. Failure in one of the member links in the Aggregated Link will not affect communication to the Distributed Fabric as traffic will be forwarded via the remaining member links. The Switches in the Distributed Fabric can support up to 4 member links in each of the 13 supported Aggregated Links.

For more information on Aggregated Links and LACP, refer to "Aggregated Links" *on* page 25.

A key feature of DLA is the ability to provide intelligent local forwarding within the Distributed Fabric, as described below.

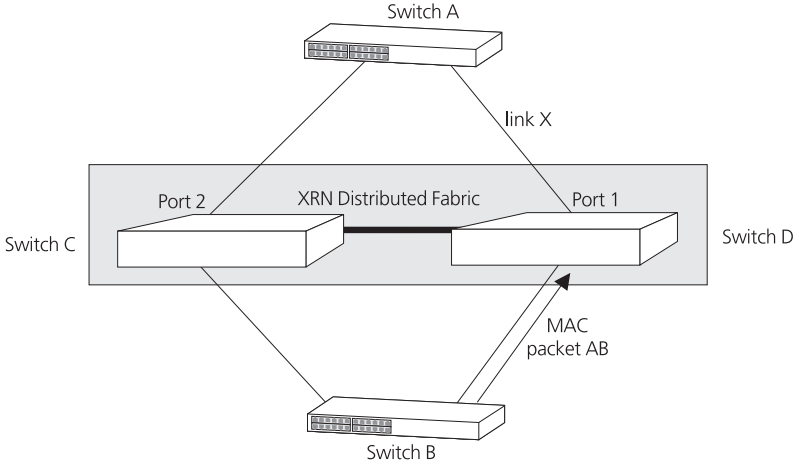
Intelligent Local Forwarding

Each Switch in the XRN Distributed Fabric can provide Intelligent Local Forwarding — this means that Layer 2 traffic originating from the edge of the network is switched locally, if it is destined for a device directly connected to one of the Switches in the Distributed Fabric. This implementation minimises the amount of traffic flowing across the XRN interconnect and therefore enhances network performance.

For example, intelligent local forwarding means that when MAC packet AB arrives at Switch D for onward transmission to Switch A, Switch D as the local Switch will use an aggregated link port selection mechanism to ensure that the packet will be forwarded to a port on the local unit, port 1 in this case. This ensures that the packet does not cross the XRN Interconnect unnecessarily.

However, if link X fails, the local forwarding rules will not apply and the ordinary aggregated link rules will ensure that the packet crosses the Interconnect and is transmitted on port 2 of Switch C to ensure it reaches its destination (Switch A).

Figure 37 Intelligent local forwarding Example



Distributed Link Aggregation Example

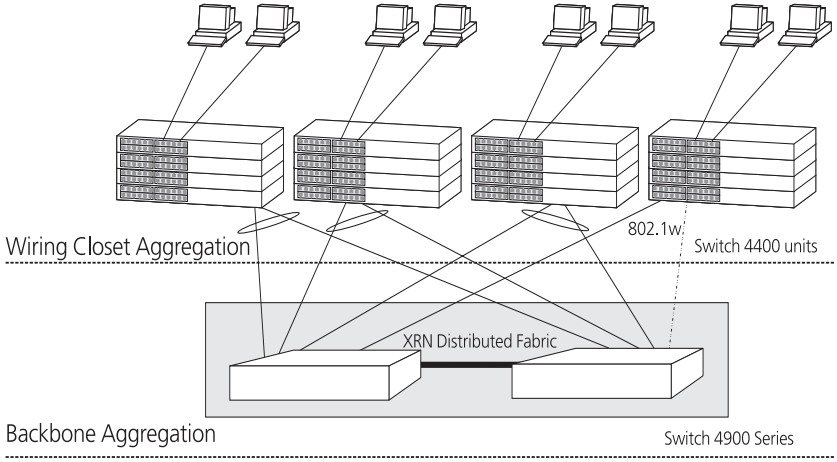
You can also use DLA to create highly resilient network backbones, supporting multihomed links to the wiring closets as shown in Figure 38.

Intelligent local forwarding ensures that each Switch in the XRN Distributed Fabric forwards traffic to local Link Aggregation ports rather than across the XRN Interconnect, thereby reducing network traffic.

You can also use resilient links or STP/RSTP for resilience, however, this does not provide the bandwidth advantage of link aggregation.

For more information about STP/RSTP refer to Chapter 4 "Using Resilience Features".

Figure 38 Distributed Link Aggregation at the Network Backbone



How to Implement XRN Technology — Overview

This section provides an overview on how to implement XRN Technology in your network. Following the steps below will ensure that your XRN network operates correctly.

1. Design your network using XRN Distributed Fabrics, taking into account all the important considerations and recommendations (see "Important Considerations and Recommendations" on page 134).
2. Ensure that the Switches containing the XRN Interconnect Module that you plan to interconnect are running the correct software (that is, version 3.0 or later).
3. Install an Interconnect Module into two supported Switches and connect with an Interconnect Cable.

   Once two Switches are interconnected in this way they create a Distributed Fabric, that is they behave as if they were one Switch and can be managed via a single IP address.
4. Set up the IP information so you can begin managing and configuring the Switches in the Distributed Fabric.

For more information on setting up IP information for your Switch so it is ready for management, refer to Chapter 3 of the Getting Started Guide that accompanies your Switch.

5 If VLANs are required (for example, if the network is in a Layer 3 environment), create the VLANs and assign VLAN membership to all ports. For more information on VLANs, see Chapter 8 "Setting Up Virtual LANs", and "VLAN-based Routing" on page 99.

6 Configure the Aggregated Links either manually or automatically via LACP, ensuring they are tagged members of all VLANs. For more information on Aggregated Links, see "Aggregated Links" on page 25.

7 Configure the router IP interfaces. For more information on router interfaces, see "Router Interfaces" on page 96 and "Establishing IP Interfaces" on page 101.

For detailed descriptions of the web interface operations and the command line interface (CLI) commands that you require to manage the Switch please refer to the Management Interface Reference Guide supplied in HTML format on the CD-ROM that accompanies your Switch or on the 3Com Web site.

Important Considerations and Recommendations

This section contains important points and recommendations that you need to consider or be aware of when designing a network using XRN technology.

- XRN support is provided by the Interconnect module and cable, therefore the Switches in the Distributed Fabric are limited to the same geographical location. Should you wish to connect two Switch units together that are located in separate racks you can obtain a 5 m Interconnect cable (3C17722). Contact your network supplier for further information.
- Bandwidth across the XRN Interconnect module is 8 Gbps. Intelligent Local Forwarding and DRR ensures efficient use of bandwidth across the XRN Interconnect.
- A maximum of 30 port-based VLANs are supported across the Distributed Fabric.
- The XRN Distributed Fabric is currently limited to two units.
- When you create a Distributed Fabric the relevant port-based tables do not double in size, they remain as they were.

Recommendations for Achieving Maximum Resilience

To achieve maximum network-level resilience 3Com recommends that:

- Servers and wiring closets are multihomed. That is, each server or wiring closet is connected to both units within the Distributed Fabric.
- On all multi-homed links you use link aggregation (preferably configured automatically via LACP rather than configured manually) across an XRN Distributed Fabric, and you have STP/RSTP enabled across your network. (See "Legacy Aggregated Links" on page 140 for more information.)

  If you are unable to use link aggregation on multihomed links, then STP/RSTP should be used as the second option, and the last option would be to use resilient links.

  This implementation increases the level of fault tolerance as it also protects against loss of the physical interfaces at the core.
- If you use the 3Com Switch 4050/4060 in a Distributed Fabric further resilience is achieved as each interface module has its own redundant power source.
- You always have STP/RSTP enabled on your network to prevent the risk of loops occuring if you have links that are multihomed, particularly if you are using link aggregation.
- All multihomed links and alternate paths must carry *all* VLANs, and packets must be tagged.
- The Distributed Fabric is the STP root bridge.
- Individual port members of each aggregated link must have VLAN membership manually configured before the aggregated link is set up. You must not rely on port members inheriting VLAN membership configuration from the aggregated link. (See "VLANs" on page 139 for more information.)
- If you are using resilient links, these must be configured on the remote unit, not on the units within the Distributed Fabric.

If you follow the 3Com recommendations, should there be a unit or interconnect failure within the Distributed Fabric, traffic flow will be maintained at all times. If you want to know more detail about how the Distributed Fabric behaves in certain failure scenarios, see "How a Failure affects the Distributed Fabric" *on* page 143.

Mixed Distributed Fabric Considerations

This section details points to consider when planning to interconnect different Switches from the SuperStack 4900 Series family or the 3Com Switch 4050/4060 family.

- If interconnecting a SuperStack 3 Switch 4924/4950 or 3Com Switch 4050/4060 with a SuperStack 3 Switch 4900/4900SX, QoS support across the Distributed Fabric will be limited to that supported by the SuperStack 3 Switch 4900/4900SX. For information on the QoS support offered by the SuperStack 3 Switch 4900/4900SX, see Chapter 6 "Using Traffic Prioritization".
- If interconnecting a Switch 4924/4950 or 3Com Switch 4050/4060 with a 4900/4900SX, Webcache support (Layer 4 Redirection) across the Distributed Fabric will not be supported.
- Single step Switch software upgrades across the Distributed Fabric can only be carried out if the Switches are of the same family.

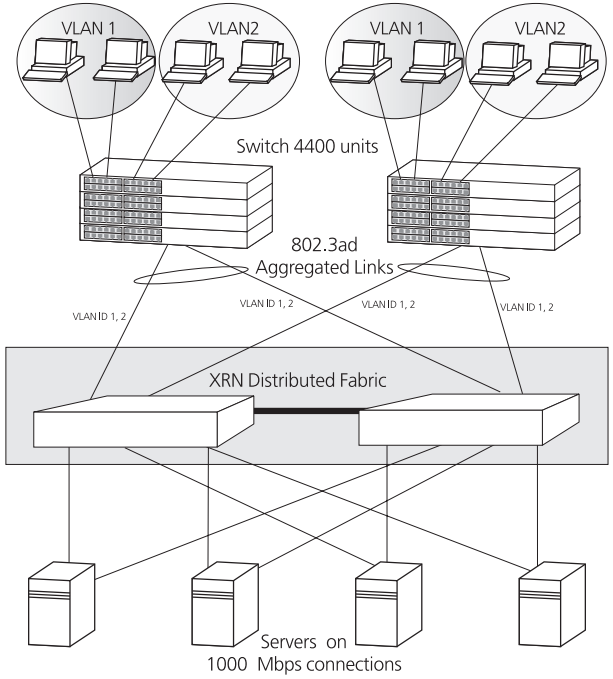
Network Example using XRN

The following example explains how to set up XRN technology in a network with a single XRN Distributed Fabric network. The same process scales for larger networks if you are using multiple XRN Distributed Fabrics.

Single XRN Distributed Fabric Network

The example in Figure 39 shows a network with two Switches interconnected to create a single XRN Distributed Fabric. The servers are multi-homed as are the Switch 4400 stacks to create a highly resilient network.

Figure 39 A single XRN Distributed Fabric Network



How to Set up this Network

This section provides information on how to configure an XRN network as shown in Figure 39. It assumes you have carried out steps 1 to 4 as detailed in "How to Implement XRN Technology — Overview" on page 133.

1 Enable LACP on the required ports, ensuring you have not connected your devices to the Distributed Fabric yet as you must configure your VLANs before the aggregated links are configured. Individual port members of each aggregated link must have VLAN membership explicitly set, that is, manually configured, before the aggregated link is set up. You must not rely on port members inheriting VLAN membership configuration from the aggregated link.

2 Create the VLANs and assign VLAN membership to all ports. For more information on VLANs, see Chapter 8 "Setting Up Virtual LANs", and "VLAN-based Routing" on page 99.

3 Connect up your ports. As LACP was enabled in step 1 the aggregated links will now automatically configure themselves. For more information on Aggregated Links, see "Aggregated Links" on page 25.

4 Configure the router IP interfaces. For more information on router interfaces, see "Router Interfaces" on page 96 and "Establishing IP Interfaces" on page 101.

5 Ensure that RSTP is enabled across the network.

Legacy aggregated links are not resilient to an interconnect failure. Hence the 3Com recommendation to use IEEE 802.3ad aggregated links (LACP) for maximum resilience.

If an automatic aggregated link (created by LACP) contains ports with different VLAN membership, the aggregated link will inherit the VLAN memberhip of the first port that comes up in the aggregated link. It will override any pre-defined VLAN membership for the aggregated link.

Recovering your XRN Network

In the event of a failure within your XRN network, 3Com recommends that you follow the recommendations below.

Unit Failure

The steps below outline the procedure to recover your XRN network in the event of a unit failure within your Distributed Fabric.

1 Obtain a Switch and ensure it is installed with same software version as the failed Switch.

2 Initialise the new Switch so it is operating with its factory default settings.

3 Connect the new Switch to the operational Switch to form the Distributed Fabric.

4 IP interfaces and VLANs will be converged between the two Switches, that is, IP interfaces and the creation of the VLANs is done automatically on the new Switch. However, any port-based configuration must be done manually.

5 If any Switch features, for example, STP or IGMP snooping are not set to default state then these should be reset after the new Distributed Fabric has been formed.

Interconnect Failure

The steps below outline the procedure to recover your XRN network in the event of an interconnect failure within your Distributed Fabric.

1 Obtain a new cable or module (depending on which item has failed).

2 Power down the two Switches.

3 Install the new cable or module.

4 Power up the Switches.

How XRN Technology Interacts with other Features

This section provides supplementary information on how the XRN Technology interacts with other software features supported by your Switch.

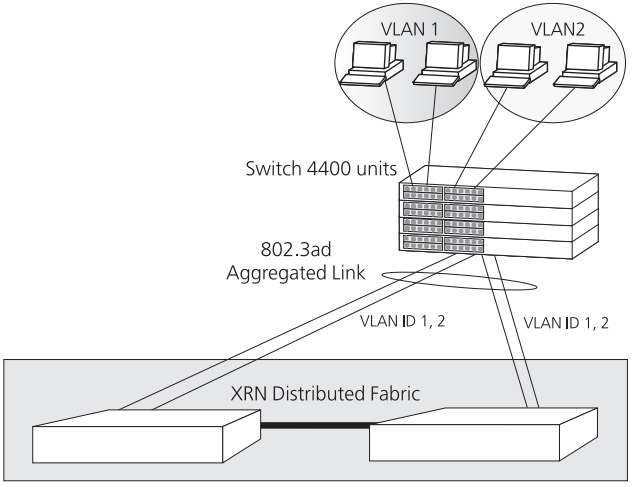
VLANs

Figure 40 shows a single aggregated link, created automatically via LACP, connecting the Switch 4400 stack to the Distributed Fabric. The Distributed Fabric will take its VLAN membership from a port within the Switch 4400 stack.

If the XRN Interconnect fails the aggregated link will split, creating two separate aggregated links (as shown in Figure 41). If the ports within the Switch 4400 stack each have different VLAN membership, this will mean the two newly formed aggregated links will also have different VLAN membership. This will result in the different VLANs not being able to communicate.

3Com recommends that you set individual ports that are to be members of an aggregated link to the same VLAN membership. This ensures communication between all VLANs at all times.

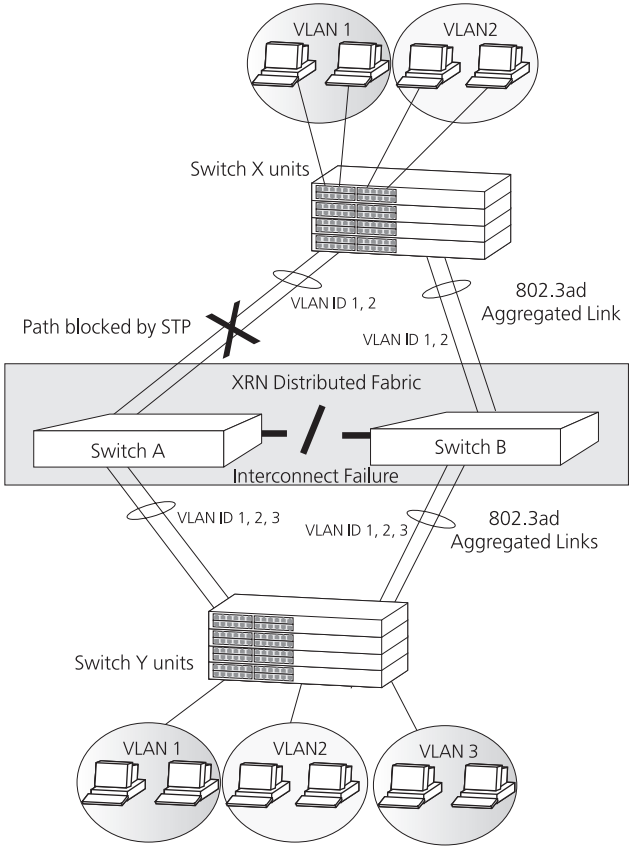
Figure 40 How XRN Technology interacts with VLANs — Example 1



The Distributed Resilient Routing (DRR) feature also requires that all units can communicate with each other on all VLANs. This ensures that on an interconnect failure all units can communicate with each other.

For example in Figure 41 the interconnect has failed and only one of the units in the Distributed Fabric will act as the router. STP will detect a potential loop and block a path of its choosing, in this example it has blocked the path between Switch X units and Switch A. If ports have different VLAN membership as shown here there will be loss of communication between VLANs 1 and 2.

Figure 41 How XRN Technology interacts with VLANs — Example 2



Legacy Aggregated Links

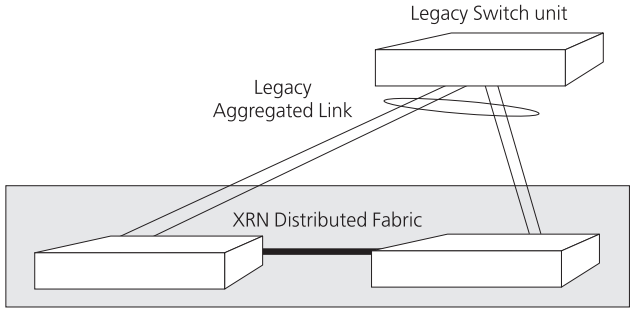
Legacy aggregated links, will react in the normal way if a unit within the Distributed Fabric fails, that is, all traffic will be redirected down the link(s) to the unit that is still operating.

However, in Figure 42, if the interconnect fails, the aggregation is still a single logical entity at the legacy Switch end, but it is now split over both

units within the Distributed Fabric. The legacy Switch is not aware that the aggregation has split and will continue to send traffic over both links, resulting in data loss.

Hence the recommendation to use IEEE 802.3ad aggregated links, if possible, as legacy aggregated links are not resilient to an XRN interconnect failure.

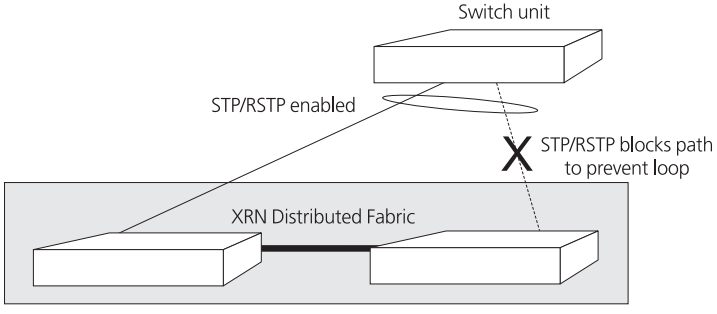
Figure 42 How XRN Technology interacts with legacy aggregated links



STP/RSTP STP/RSTP should be used for multihomed links if you are not able to use aggregated links. Figure 43 shows how STP will prevent a loop occuring on a multihomed link.

STP/RSTP should always be enabled if your multihomed links are aggregated links. Figure 41 shows how, on interconnect failure, STP/RSTP will detect the potential loop caused by the aggregated links splitting and block a path to prevent the loop occuring.

Figure 43 How XRN Technology interacts with STP/RSTP



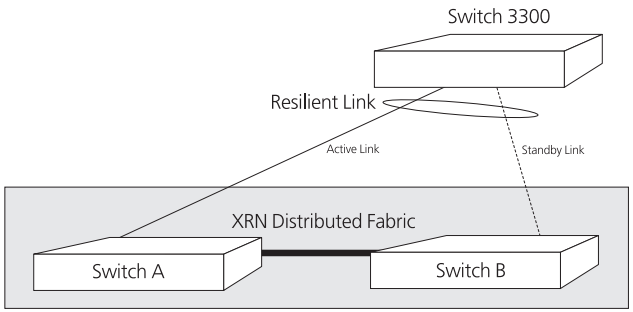
Resilient Links

In Figure 44, if Switch A within the Distributed Fabric fails, the Switch 3300 will detect that a link has gone down and will make the standby link to Switch B active and pass all traffic down the link to Switch B.

When using resilient links in a Distributed Fabric network the resilient links must be configured at the remote end rather than at the Distributed Fabric. In a unit failure scenario as described above it would not matter if the resilient links were configured at the Distributed Fabric end. However, on an interconnect failure it would matter.

For example, if the resilient links were configured on Switches A and B, if the interconnect fails, both Switches will detect a failed link to Switch 3300 and both A and B will activate their links to Switch 3300. So both links in the resilient link will be passing traffic, potentially causing a loop in the network.

Figure 44 How XRN Technology interacts with Resilient Links



How a Failure affects the Distributed Fabric

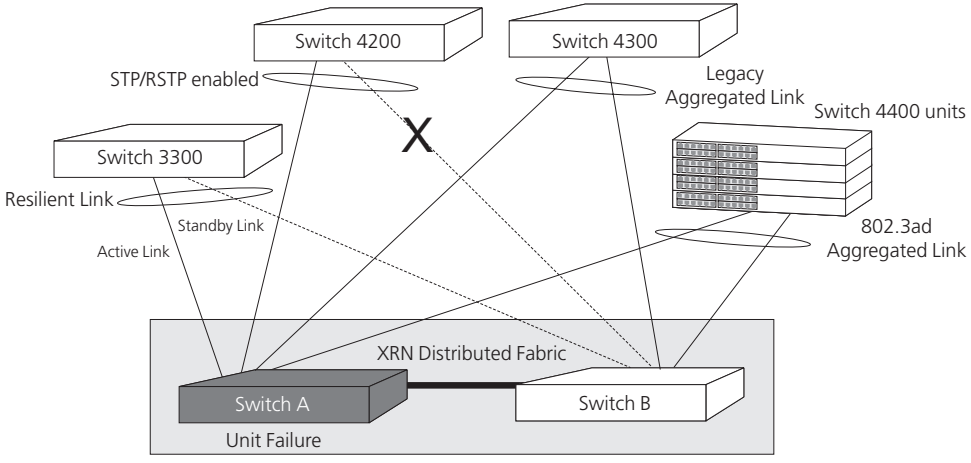
This section provides supplementary information on how the Distributed Fabric and traffic flow is affected by failure of an XRN Interconnect and of a unit in the Distributed Fabric.

Loss of a Switch within the XRN Distributed Fabric

When a Switch unit in the Distributed Fabric fails, assuming you have followed the recommendations in "Important Considerations and Recommendations" on page 134, your traffic flow should continue through your network.

The way the network reacts depends upon which features are configured on which links. For example, Figure 45 shows an XRN network where all the edge devices are connected to the Distributed Fabric using a range of supported features, some of which are legacy features.

Figure 45 XRN Network reaction on Distributed Fabric unit failure



Should Switch A fail, the network will react in the following way:

LACP (IEEE 802.3ad) and Legacy Aggregated Links

The Switch 4400 and Switch 4300 Aggregated Links will reroute all traffic down the link connected to Switch B.

Legacy STP (IEEE802.1D) and RSTP (IEEE 802.1w)

The Switch 4200 is using legacy STP. STP will reconfigure the network to open the previously blocked link to Switch B. The STP reconfiguration will cause all Switch forwarding databases (MAC address tables) to be fast aged (if using RSTP, they will be flushed).

Resilient Links

The Switch 3300 is using resilient links, which should be set up at the 3300 end of the link. The Switch 3300 will detect the unit failure and activate the link to Switch B

VLANs

Any VLANs will not be affected by unit failure.

Router

Switch B will continue to do all the routing. As it was routing prior to Switch A's failure there will be no change of the router identity, that is, the router interface IP addresses will not change. The router interface MAC addresses may change but this will have no visible impact on your network. Any MAC address change is propogated to your network by the issuing of gratuitous ARP messages.

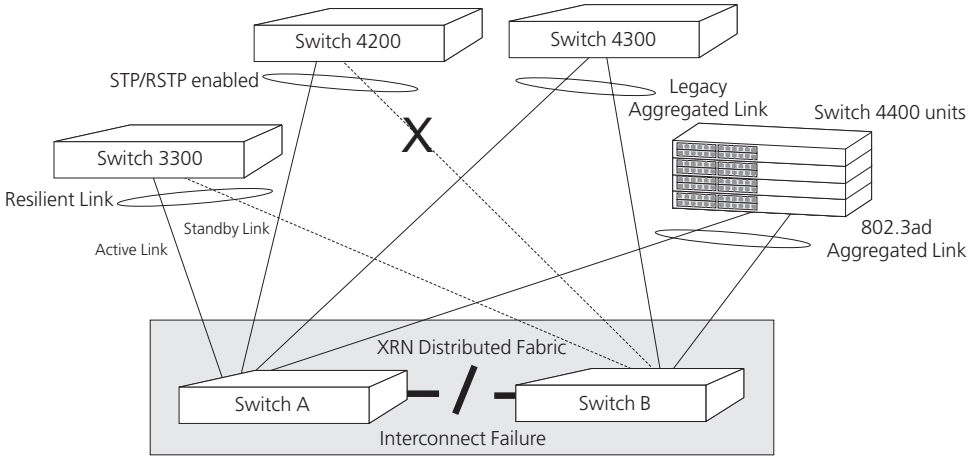
Switch A Recovery

When Switch A recovers and starts to operate again, all links will reconfigure themselves as they were before the failure, according to the protocols used. The routing task will once again be shared between Switches A and B, using the same IP address.

Loss of the XRN Interconnect

When an interconnect fails between two Switches in the Distributed Fabric, assuming you have STP/RSTP and LACP enabled as recommended in "Important Considerations and Recommendations" on page 134, your traffic flow should continue through your network.

Figure 46 XRN Network reaction on XRN Interconnect failure



In Figure 46, if the interconnect fails, the network will react in the following way:

LACP (IEEE 802.3ad) and Legacy Aggregated Links

The Switch 4400 automatically configured aggregated link (LACP) will reconfigure itself to create two separate aggregated links.

The Switch 4300 legacy aggregated link will be split between the two Switches in the Distributed Fabric and will no longer operate and will cause network disruption.

Legacy aggregated links are not resilient to an interconnect failure. Hence the 3Com recommendation to use IEEE 802.3ad aggregated links (LACP) for maximum resilience.

IEEE802.1D (Legacy STP) and RSTP

The Switch 4200 is using legacy STP. STP (and RSTP) will reconfigure the network to open the previously blocked link to Switch B. The STP reconfiguration will cause all Switch forwarding databases (MAC address tables) to be fast aged (if using RSTP, they will be flushed). If STP is enabled throughout the network, it will reconfigure the network to ensure that no loops occur due to split aggregated links.

If the Distributed Fabric has been configured to be the root bridge in the network then this will encourage STP to maintain the traffic flow through the shortest paths in the event of an XRN Interconnect failure.

Resilient Links

The Switch 3300 will continue to send traffic down the active link to Switch A and keep the link to Switch B in standby mode.

VLANs

As all VLANs will have been configured on all links, the traffic will still reach its destination via the paths that remain open.

Router

Initially both Switches continue to route. Simultaneously the Switches attempt to contact each other and carry out a process that determines which Switch will become the Layer 3 router and which Switch will become the Layer 2 bridge. This avoids the scenario of two different and independent routers operating with the same identity.

Interconnect Recovery

When the interconnect recovers and starts to operate again, all links will reconfigure themselves as they were before the failure, according to the protocols used. The routing task will once again be shared between Switches A and B, using the same IP address.

A CONFIGURATION RULES

Configuration Rules for Gigabit Ethernet

Gigabit Ethernet is designed to run over several media:

- Single-mode fiber optic cable, with connections up to 5 km (3.1 miles). Support for distances over 5 km is supported depending on the module specification.
- Multimode fiber optic cable, with connections up to 550 m (1804 ft).
- Category 5 cabling, with connections up to 100 m (328 ft).

The different types of Gigabit Ethernet media and their specifications are detailed in Table 16.

Table 16 Gigabit Ethernet cabling

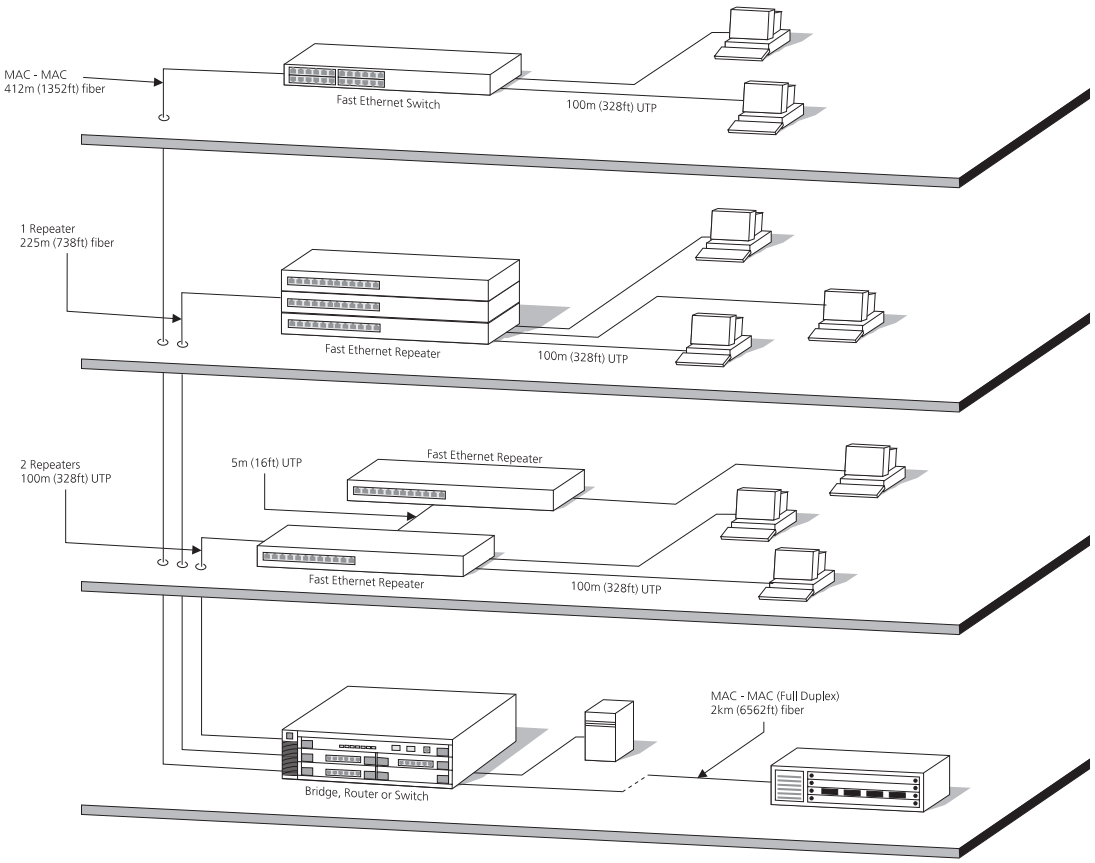
Gigabit Ethernet Transceivers	Fiber Type	Modal Bandwidth (MHz/km)	Lengths Supported Specified by IEEE (meters)
1000BASE-LX	62.5 µm MM	500	2–550
	50 µm MM	400	2–550
	50 µm MM	500	2–550
	10 µm SM	N/A	2–5000
1000BASE-SX	62.5 µm MM	160	2–220
	62.5 µm MM	120	2–275
	50 µm MM	400	2–500
	50 µm MM	500	2–550
1000BASE-T	N/A	N/A	100

MM = Multimode *SM = Single-mode*

Configuration Rules for Fast Ethernet

The topology rules for 100 Mbps Fast Ethernet are slightly different to those for 10 Mbps Ethernet. Figure 47 illustrates the key topology rules and provides examples of how they allow for large-scale Fast Ethernet networks.

Figure 47 Fast Ethernet configuration rules



The key topology rules are:

- Maximum UTP cable length is 100 m (328 ft) over Category 5 cable.
- A 412 m (1352 ft) fiber link is allowed for connecting switch-to-switch, or endstation-to-switch, using half-duplex 100BASE-FX.
- A total network span of 325 m (1066 ft) is allowed in single-repeater topologies (one hub stack per wiring closet with a fiber link to the

collapsed backbone). For example, a 225 m (738 ft) fiber link from a repeater to a router or switch, plus a 100 m (328 ft) UTP link from a repeater out to the endstations.

Configuration Rules with Full Duplex

The Switch provides full duplex support for all its ports, including Expansion Module ports. Full duplex allows packets to be transmitted and received simultaneously and, in effect, doubles the potential throughput of a link.

With full duplex, the Ethernet topology rules are the same, but the Fast Ethernet rules are:

- Maximum UTP cable length is 100 m (328 ft) over Category 5 cable.
- A 2 km (6562 ft) fiber link is allowed for connecting switch-to-switch, or endstation-to-switch.

B NETWORK CONFIGURATION EXAMPLES

This chapter contains the following sections:

The illustrations depict a 4900 Series device. These configurations apply to all Switches in the Switch 4900 Series.

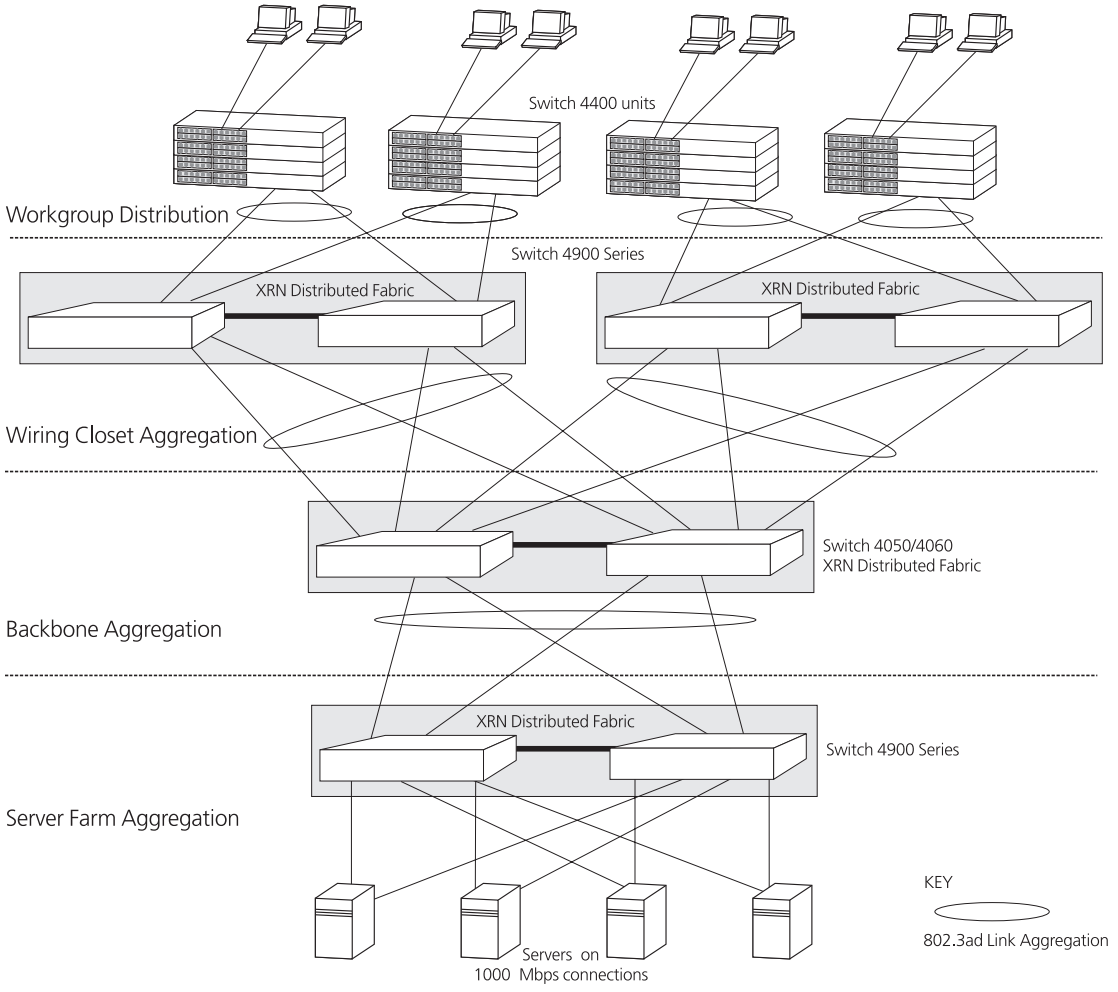
Network Configuration Examples

This section shows some network examples that illustrate how you can set up your network for optimum performance using some of the features supported by your Switch.

Maximizing the Resilience of Your Network

Figure 48 shows how you can set up your network to maximize its resilience using XRN Technology. All devices are multihomed, aggregated links are created using LACP, and RSTP is enabled to prevent loops occuring in the event of a link failure.

Figure 48 Network set up to maximize resilience

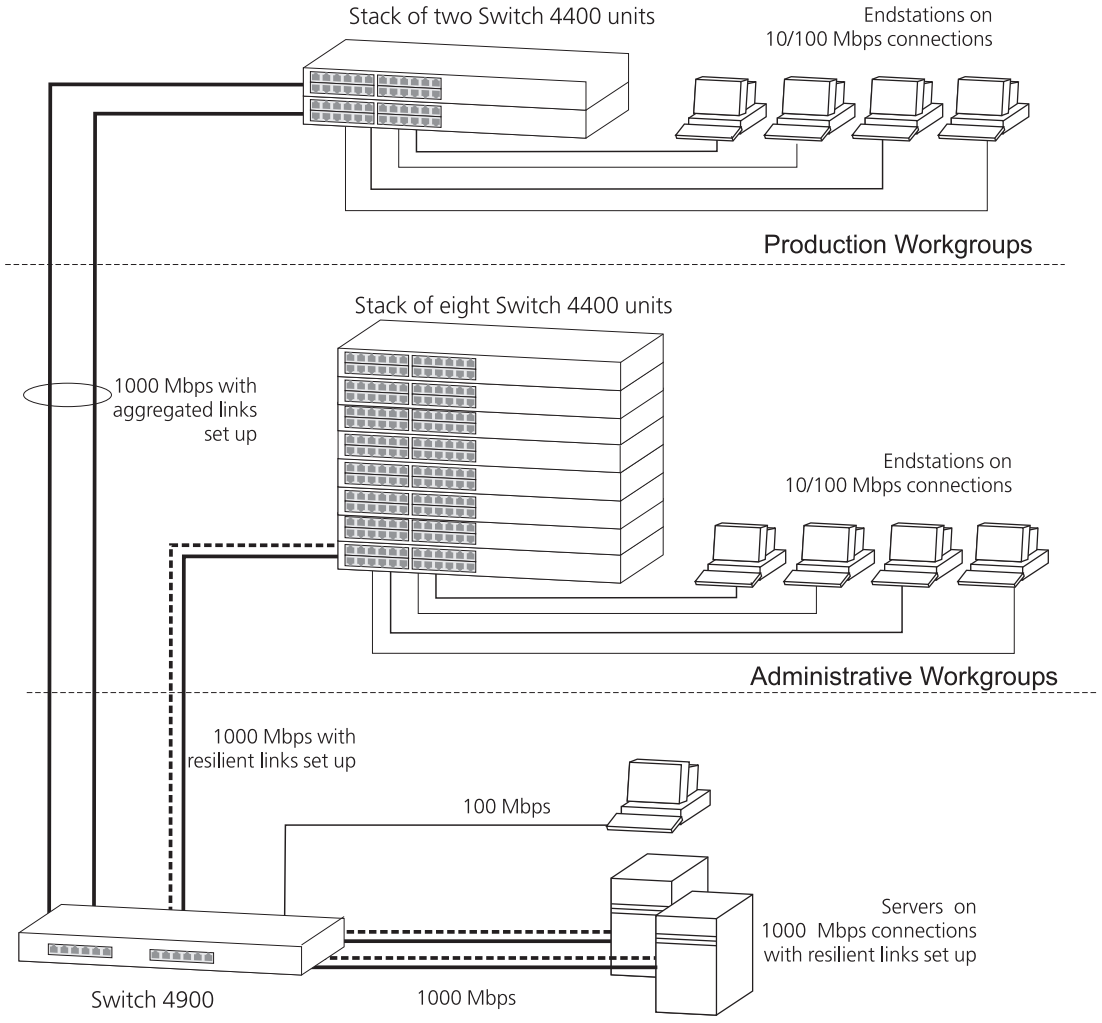


Enhancing the Performance of Your Network

Figure 49 shows how you can set your network up to enhance its performance.

All ports are auto-negotiating and smart auto-sensing and will therefore pass data across the network at the optimum available speed and duplex mode. Flow control will help avoid packet loss during periods of network congestion. A Gigabit Ethernet backbone is set up between the Switch 4900 and each Switch in the workgroups to increase the bandwidth, and therefore the overall network performance.

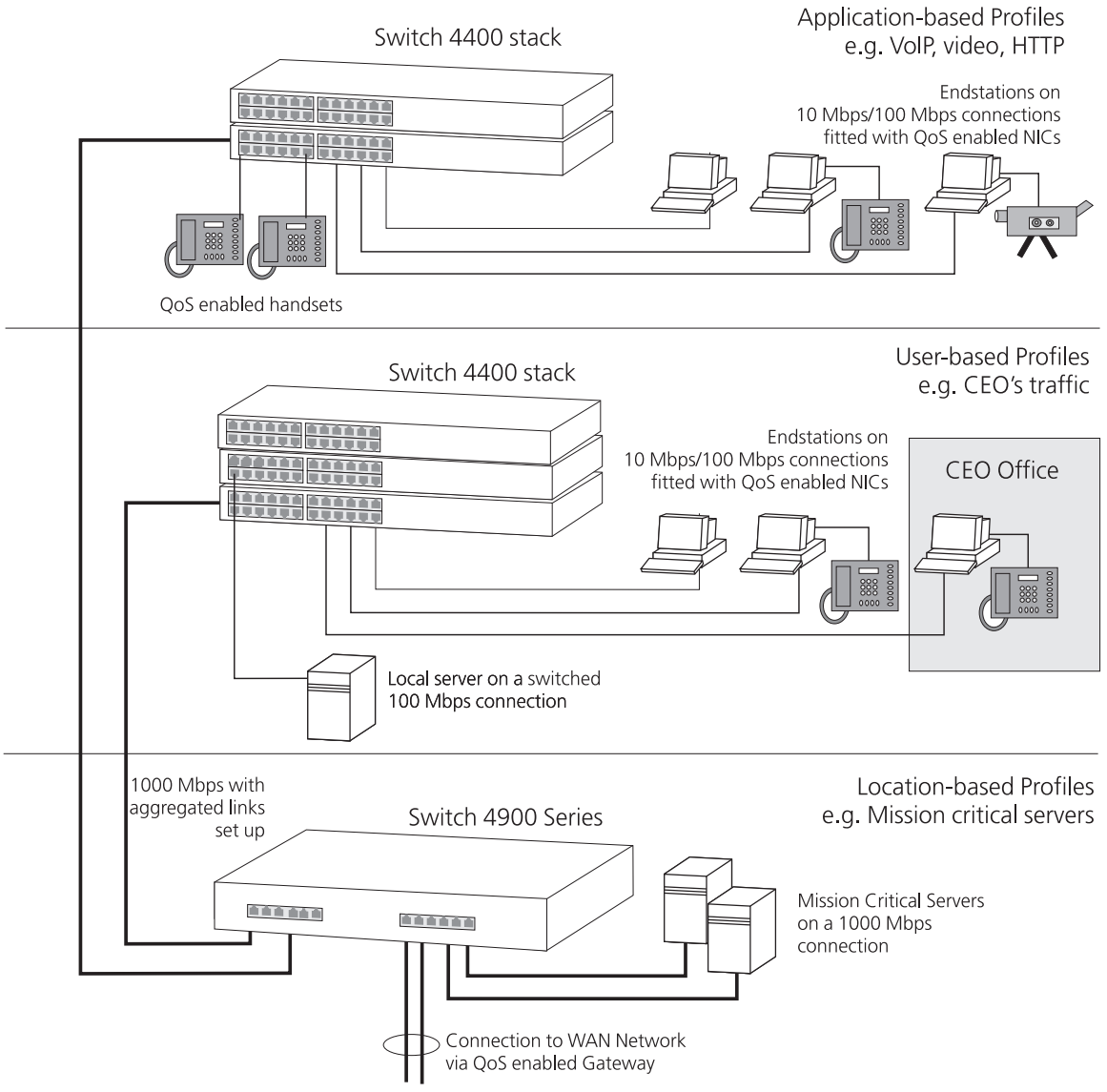
Figure 49 Network set up to enhance performance



Utilizing the Traffic Prioritization Features of Your Network

The example in Figure 50 shows a network configuration that demonstrates how you can utilize the different types of Quality of Service (QoS profiles) to ensure a high level of service and prioritization across the network for certain applications, users, or locations. For more information on using QoS, see Chapter 6 "Using Traffic Prioritization".

Figure 50 Network set up to utilize traffic prioritization



C IP ADDRESSING

This chapter provides some background detail on the IP information that needs to be assigned to your Switch to enable you to manage it across a network. The topics covered are:

- IP Addresses
- Subnets and Subnet Masks
- Default Gateways
- Standards, Protocols, and Related Reading

IP addressing is a vast topic and there are white papers on the World Wide Web and publications available if you wish to learn more about IP addressing.

IP Addresses

This IP address section is divided into two parts:

- Simple Overview — Gives a brief overview of what an IP address is.
- Advanced Overview — Gives a more in depth explanation of IP addresses and the way they are structured.

Simple Overview

To operate correctly, each device on your network must have a unique IP address. IP addresses have the format *n.n.n.n* where *n* is a decimal number between 0 and 255. An example IP address is '192.168.100.8'.

The IP address can be split into two parts:

- The first part, called the network part, ('192.168' in the example) identifies the network on which the device resides.
- The second part, called the host part, ('100.8' in the example) identifies the device within the network.

If your network is internal to your organization only, you may use any arbitrary IP address. 3Com suggests you use addresses in the series 192.168.100.*X* (where *X* is a number between 1 and 254) with a subnet mask 255.255.255.0. If you are using SLIP, use the default SLIP address of 192.168.101.1 with a subnet mask of 255.255.255.0.

These suggested IP addresses are part of a group of IP addresses that have been set aside specially for use "in house" only.

CAUTION: *If your network has a connection to the external IP network, you must apply for a registered IP address. This registration system ensures that every IP address used is unique; if you do not have a registered IP address, you may be using an identical address to someone else and your network will not operate correctly.*

Obtaining a Registered IP Address

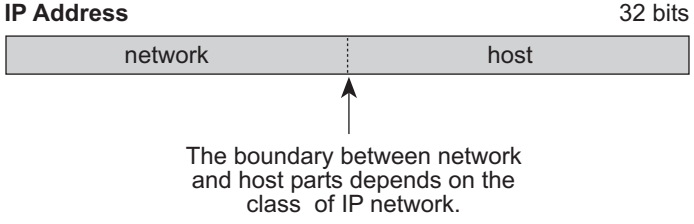
InterNIC Registration Services is the organization responsible for supplying registered IP addresses. The following contact information is correct at time of publication:

World Wide Web site: **`http://www.internic.net`**

Advanced Overview

IP addresses are 32-bit addresses that consist of a *network part* (the address of the network where the host is located) and a *host part* (the address of the host on that network).

Figure 51 IP Address: Network Part and Host Part



IP addresses differ from Ethernet MAC addresses, which are unique hardware-configured 48-bit addresses. A central agency, such as the InterNIC Registration Services mentioned above, assigns the network part of the IP address, and you assign the host part. All devices that are connected to the same network share the same network part (also called the *prefix*).

Dotted Decimal Notation

The actual IP address is a 32-bit number that is stored in binary format. These 32 bits are segmented into 4 groups of 8 bits — each group is referred to as a *field* or an *octet*. Decimal notation converts the value of each field into a decimal number, and the fields are separated by dots.

Figure 52 Dotted Decimal Notation for IP Addresses

```
10011110.01100101.00001010.00100000     = Binary notation

                    158.101.10.32       = Decimal notation
```

The decimal value of an octet whose bits are all 1s is 255.

Network Portion

The location of the boundary between the network part and the host part depends on the class that the central agency assigns to your network. The three primary classes of IP addresses are as follows:

- **Class A address** — Uses 8 bits for the network part and 24 bits for the host part. Although only a few Class A networks can be created, each can contain a very large number of hosts.
- **Class B address** — Uses 16 bits for the network part and 16 bits for the host part.
- **Class C address** — Uses 24 bits for the network part and 8 bits for the host part. Each Class C network can contain only 254 hosts, but many such networks can be created.

The high-order bits of the network part of the address designate the IP network class. See Table 17.

Table 17 How Address Class Corresponds to the Address Number

Address Class	High-order Bits	Address Number (Decimal)
A	0nnnnnnn	0-127
B	10nnnnnn	128-191
C	11nnnnnn	192-254

Subnets and Subnet Masks

You can divide your IP network into sub-networks also known as subnets. Support for subnets is important because the number of bits assigned to the device part of an IP address limits the number of devices that may be addressed on any given network. For example, a Class C address is restricted to 254 devices.

The IP address can also contain a *subnetwork part* at the beginning of the host part of the IP address. Thus, you can divide a single Class A, B, or C network internally, allowing the network to appear as a single network to other external networks. The subnetwork part of the IP address is visible only to hosts and gateways on the subnetwork.

When an IP address contains a subnetwork part, a *subnet mask* identifies the bits that constitute the subnetwork address and the bits that constitute the host address. A subnet mask is a 32-bit number in the IP address format. The *1* bits in the subnet mask indicate the network and subnetwork part of the address. The *0* bits in the subnet mask indicate the host part of the IP address, as shown in Figure 53.

Figure 53 Subnet Masking

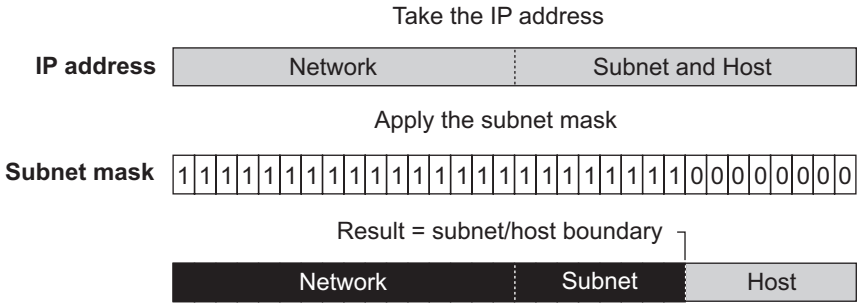


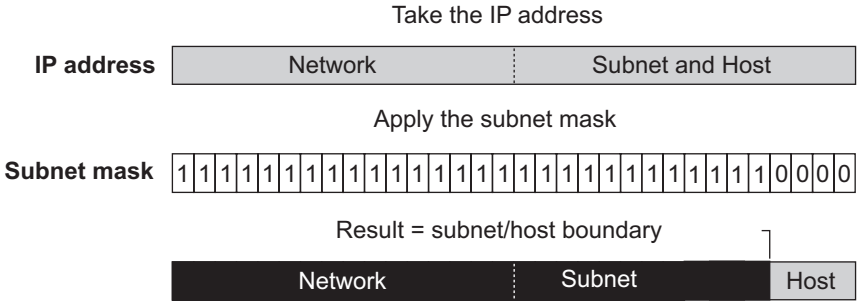
Figure 54 shows an example of an IP address that includes network, subnetwork, and host parts. Suppose the IP address is *158.101.230.52* with a subnet mask of *255.255.255.0*. Since this is a Class B address, this address is divided as follows:

- *158.101* is the network part
- *230* is the subnetwork part
- *52* is the host part

As shown in this example, the 32 bits of an IP address and subnet mask are usually written using an integer shorthand. This notation translates four consecutive 8-bit groups (octets) into four integers that range from 0 through 255. The subnet mask in the example is written as 255.255.255.0.

Traditionally, subnet masks were applied to octets in their entirety. However, one octet in the subnet mask can be further subdivided so that part of the octet indicates an *extension* of the network number, and the rest of the same octet indicates the host number, as shown in Figure 54.

Figure 54 Extending the Network Prefix



Using the Class B IP address from Figure 53 (158.101.230.52), the subnet mask is 255.255.255.240.

The number that includes both the Class B natural network mask (255.255) and the subnet mask (255.240) is sometimes called the *extended network prefix*.

Continuing with the previous example, the subnetwork part of the mask uses 12 bits, and the host part uses the remaining 4 bits. Because the octets are actually binary numbers, the number of subnetworks that are possible with this mask is 4,096 (2^{12}), and the number of hosts that are possible in each subnetwork is 16 (2^4).

Subnet Mask Numbering

An alternate method to represent the subnet mask numbers is based on the number of bits that signify the network portion of the mask. Many Internet Service Providers (ISPs) now use this notation to denote the subnet mask. See Table 18.

Table 18 Subnet Mask Notation

Standard Mask Notation	Network Prefix Notation
100.100.100.100 (255.0.0.0)	100.100.100.100/8
100.100.100.100 (255.255.0.0)	100.100.100.100/16
100.100.100.100 (255.255.255.0)	100.100.100.100/24

The subnet mask 255.255.255.255 is reserved as the default broadcast address.

Default Gateways

A gateway is a device on your network which is used to forward IP packets to a remote destination. An alternative name for a gateway is a Router. "Remote" refers to a destination device that is not directly attached to the same network segment as the source device.

The source device cannot send IP packets directly to the destination device because it is in a different network segment. Instead you configure it to send the packets to a gateway which is attached to multiple segments.

When it receives the IP packets, the gateway determines the next network hop on the path to the remote destination, and sends the packets to that hop. This could either be the remote destination or another gateway closer towards the destination.

This hop-by-hop process continues until the IP packets reach the remote destination.

If manually configuring IP information for the Switch, enter the IP address of the default gateway on the local subnet in which the Switch is located. If no default gateway exists on your network, enter the IP address `0.0.0.0` or leave the field blank.

Standards, Protocols, and Related Reading

This section describes how to obtain more technical information about IP.

Requests For Comments (RFCs)

Documents called Requests for Comments (RFCs) contain information about the entire set of protocols that make up IP. Some of the RFCs that pertain to the discussions in this chapter are:

- **RFC 791** — Internet Protocol
- **RFC 1219** — Subnetwork Numbers
- **RFC 1878** — VLSMs
- **RFC 1519** — Supernetting
- **RFC 1256** — ICMP Router Discovery Messages
- **RFC 1058** — RIP
- **RFC 1723** — RIP Version 2
- **RFC 1786** — IP Routing Policies
- **RFC 2400** — Internet Official Protocol Standards

You can obtain RFCs from the Internet using the following URL:

`http://sunsite.auc.dk/RFC`

Standards Organizations

Standards organizations ensure interoperability, create reports, and recommend solutions for communications technology. The most important standards groups are:

- International Telecommunications Union (ITU)
- Electronic Industry Association (EIA)
- American National Standards Institute (ANSI)
- International Standards Organization (ISO)
- Institute of Electrical and Electronic Engineers (IEEE)
- Internet Engineering Task Force (IETF)
- National Institute of Standards and Technology (NIST)

D ADVANCED IP ROUTING CONCEPTS

This chapter provides some additional background detail on the IP information that can be assigned to your Switch to enable you to manage it across a network. These are advanced features and are not required for operating your switch in your network. The topics covered are:

- Variable Length Subnet Masks (VLSMs)
- Supernetting

Variable Length Subnet Masks (VLSMs)

With Variable Length Subnet Masks (VLSMs), each subnetwork under a network can use its own subnet mask. Therefore, with VLSM, you can get more subnetwork space out of your assigned IP address space.

How VLSMs Work

VLSMs get beyond the restriction that a single subnet mask imposes on the network. One subnet mask per IP network address fixes the number of subnetworks and the number of hosts per subnetwork.

For example, if you decide to configure the 158.100.0.0/16 network with a /23 extended-network prefix, you can create 128 subnetworks with each having up to 510 hosts. If some of the subnetworks do not need that many hosts, you would assign many host IP addresses but not use them.

With VLSMs, you can assign another subnet mask, for instance, /27, to the same IP address. So you can assign a longer subnet mask that consequently uses fewer host IP addresses. As a result, routing tables are smaller and more efficient.

This method of further subdividing addresses using VLSMs is being used increasingly more as networks grow in size and number. However, be aware that this method of addressing can greatly increase your network

maintenance and the risk of creating erroneous addresses unless you plan the addressing scheme properly.

Guidelines for Using VLSMs

Consider the following guidelines when you implement VLSMs:

- When you design the subnetwork scheme for your network, do not estimate the number of subnetworks and hosts that you need. Work from the top down until you are sure that you have accounted for all the hosts, present and future, that you need.
- Make sure that the routers forward routes based on what is known as the *longest match*.

  For example, assume that the destination IP address of a packet is 158.101.26.48 and that the following four routes are in the routing table:

  - 158.101.26.0/24
  - 158.101.3.10/16
  - 158.101.26.32/16
  - 158.95.80.0/8

  The router selects the route to 158.101.26.0/24 because its extended network prefix has the greatest number of bits that correspond to the destination IP address of the packet.

See RFCs 1219 and 1878 for information about understanding and using VLSMs.

Supernetting

Because Class B Internet addresses are in short supply, larger networks are now usually granted a contiguous block of several Class C addresses. Unfortunately, this creates very large routing tables since multiple Class C routes have to be defined for each network containing more than 254 nodes. Larger routing tables mean more work for the routers and, therefore, poorer performance.

Supernetting is only supported by RIPv2.

With traditional IP, each class C network must have a routing table entry.

Supernetting, or CIDR (Classless InterDomain Routing), is a technique that allows each of these larger networks to be represented by a single

routing table entry. (See RFC 1519 for detailed information about Supernetting.)

To do this, supernet addressing does something very different from traditional TCP/IP routing (which allows only one netmask per network). In supernet routing, each supernet can be assigned its own netmask.

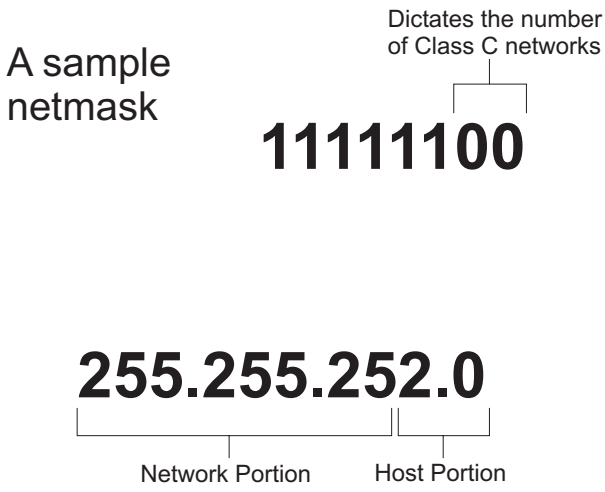
Since supernet addressing is a fairly complex mechanism, the easiest way to understand it is to step through the setup process.

Step 1 - Select a netmask for each supernet

Each supernet must have a netmask assigned to it. The netmask for an individual supernet can be, but does not have to be, the same as the netmask for any other supernet.

As in subnetting, a netmask creates a division between the network portion of an address and the host portion of an address. However, since the network you are defining is larger than a Class C network, the division you are creating is not in the fourth octet of the address. This example creates supernets composed of fewer than 254 Class C networks. So, their netmasks are actually splitting up the third octet in their IP addresses. See Figure 55.

Figure 55 Sample CIDR Netmask



Notice that the number of zero bits in the third octet actually dictates the number of Class C networks in the supernet. Each zero bit makes the

supernet twice as large. So, a supernet composed of 8 Class C networks would actually have 3 zeroes ($8 = 2^3$).

This would seem very limited since it restricts you to using groups that nicely fit into a power of 2 (1, 2, 4, 8, 16...). However, inconveniently-sized supernets can be accommodated because of a simple fact: a netmask with more 1 bits will override a netmask with fewer 1 bits.

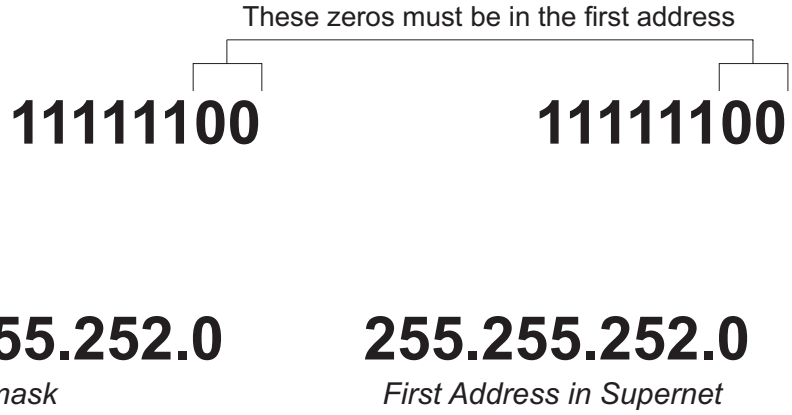
This allows a smaller supernet to share the address space of a larger supernet. If, for example, you had a supernet of size 6 and a supernet of size 2, you could assign the larger supernet an 8 network address space and assign the smaller supernet the portion of that address space that the larger supernet was not using.

Because the smaller supernet netmask has more 1 bits, packets whose address was part of its address space would be routed to the smaller supernet even though the address is also part of the address space dictated by the larger supernet netmask.

Step 2 - Select a range of addresses for each supernet

The range of addresses in a supernet must fit exactly into a space that can be described by its netmask. This means that the zero bits in the netmask must also appear in the first address of the supernet block. For this to be true, the third octet in the address must be an even multiple of the same power of 2 used to form the netmask. For example, if you had created a block of 8 networks, the third octet in the first address will be an even multiple of 8. See Figure 56.

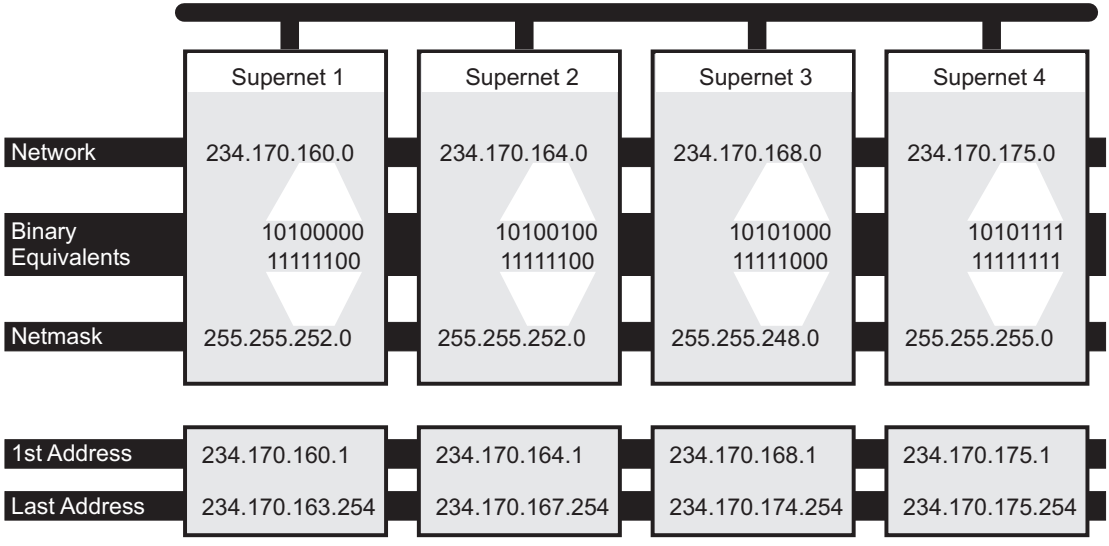
Figure 56 Selecting a Range of Addresses



Supernet Example

The four networks in Figure 57 are all connected to the same Internet service provider (ISP). The ISP has decided to use supernetting to reduce the size of the routing tables and improve throughput.

Figure 57 Supernet example



- Supernets 1 and 2 each require four Class C networks, so they require a netmask with 2 zero bits ($4 = 2^2$) in the third octet. This yields a netmask of 255.255.252.0.

- Supernet 3 requires 7 Class C address spaces. Since 7 isn't a power of 2, we have to round it up to eight. This gives it a netmask of 255.255.248.0.
- Supernet 4 is a single Class C network, making it's netmask 255.255.255.0

Now, assign ranges of addresses. Assume that the ISP is responsible for the network 234.170.0.0 and that its first free addresses are at 234.170.158.0.

The third octet of Supernet 1 has to be an even multiple of 4, so the ISP grants an address range starting at 234.170.160.0 and hopes that the block between 158 and 160 can be filled in later.

Supernet 2 must also begin on an even multiple of 4. The first available address after Supernet 1 conveniently fits the bill. So, supernet 2 extends from 234.170.164.1 to 234.170.167.254.

Supernet 3 requires an even multiple of 8. It also can begin on the next available address.

Since supernet 4 can fit entirely in a single Class C address space, it can use the supernet 3 surplus space. It is therefore given the last Class C address space in the Supernet 3 territory, effectively reducing supernet 3 to only the 7 class C networks it needs.

GLOSSARY

10BASE-T The IEEE specification for 10 Mbps Ethernet over Category 3, 4 or 5 twisted pair cable.

100BASE-FX The IEEE specification for 100 Mbps Fast Ethernet over fiber-optic cable.

100BASE-TX The IEEE specification for 100 Mbps Fast Ethernet over Category 5 twisted-pair cable.

1000BASE-T The IEEE specification for 1000 Mbps Gigabit Ethernet over four-pair Category 5 twisted-pair cable.

1000BASE-SX The IEEE specification for 1000 Mbps Gigabit Ethernet over fiber-optic cable.

aging The automatic removal of dynamic entries from the Switch Database which have timed-out and are no longer valid.

Aggregated Links Aggregated links allow a user to increase the bandwidth and resilience between switches by using a group of ports to carry traffic between the switches.

auto-negotiation A feature on twisted pair ports that allows them to advertise their capabilities for speed, duplex and flow control. When connected to a port that also supports auto-negotiation, the link can automatically configure itself to the optimum setup.

backbone The part of a network used as a primary path for transporting traffic between network segments.

bandwidth The information capacity, measured in bits per second, that a channel can transmit. The bandwidth of Ethernet is 10 Mbps, the bandwidth of Fast Ethernet is 100 Mbps, and the bandwidth of Gigabit Ethernet is 1000 Mbps.

baud The signalling rate of a line, that is, the number of transitions (voltage or frequency changes) made per second. Also known as *line speed*.

BOOTP The BOOTP protocol allows you to automatically map an IP address to a given MAC address each time a device is started. In addition, the protocol can assign the subnet mask and default gateway to a device.

bridge A device that interconnects two LANs of a different type to form a single logical network that comprises of two network segments.

Bridges learn which endstations are on which network segment by examining the source addresses of packets. They then use this information to forward packets based on their destination address. This process is known as filtering.

broadcast A packet sent to all devices on a network.

broadcast storm Multiple simultaneous broadcasts that typically absorb all the available network bandwidth and can cause a network to fail. Broadcast storms can be due to faulty network devices.

collision A term used to describe two colliding packets in an Ethernet network. Collisions are a part of normal Ethernet operation, but a sudden prolonged increase in the number of collisions can indicate a problem with a device, particularly if it is not accompanied by a general increase in traffic.

CSMA/CD Carrier-sense Multiple Access with Collision Detection. The protocol defined in Ethernet and IEEE 802.3 standards in which devices transmit only after finding a data channel clear for a period of time. When two devices transmit simultaneously, a collision occurs and the colliding devices delay their retransmissions for a random length of time.

DHCP Dynamic Host Control Protocol.

DSCP Differentiated Services Code Point. The DSCP is an identifying portion of code that determines the behavior of packets within the IP network.

endstation A computer, printer or server that is connected to a network.

Ethernet A LAN specification developed jointly by Xerox, Intel and Digital Equipment Corporation. Ethernet networks use CSMA/CD to transmit packets at a rate of 10 Mbps over a variety of cables.

Ethernet address See *MAC address*.

Fast Ethernet An Ethernet system that is designed to operate at 100Mbps.

forwarding The process of sending a packet toward its destination using a networking device.

Forwarding Database See *Switch Database*.

filtering The process of screening a packet for certain characteristics, such as source address, destination address, or protocol. Filtering is used to determine whether traffic is to be forwarded, and can also prevent unauthorized access to a network or network devices.

flow control A mechanism that prevents packet loss during periods of congestion on the network. Packet loss is caused when devices send traffic to an already overloaded port on a Switch. Flow control prevents packet loss by inhibiting devices from generating more traffic until the period of congestion ends.

full duplex A system that allows packets to be transmitted and received at the same time and, in effect, doubles the potential throughput of a link.

Gigabit Ethernet IEEE standard 802.3z for 1000 Mbps Ethernet; it is compatible with existing 10/100 Mbps Ethernet standards.

half duplex A system that allows packets to transmitted and received, but not at the same time. Contrast with *full duplex*.

hub A device that regenerates LAN traffic so that the transmission distance of that signal can be extended. Hubs are similar to repeaters, in that they connect LANs of the same type; however they connect more LANs than a repeater and are generally more sophisticated.

IEEE Institute of Electrical and Electronics Engineers. This American organization was founded in 1963 and sets standards for computers and communications.

IEEE Std 802.1D, 1998 Edition A standard that defines the behavior of bridges in an Ethernet network.

IEEE Std 802.1p A standard that defines traffic prioritization. 802.1p is now incorporated into the relevant sections of the IEEE Std 802.1D, 1998 Edition.

IEEE Std 802.1Q-1998 A standard that defines VLAN tagging.

IEEE Std 802.3ad A standard that defines link aggregation. 802.3ad is now incorporated into the relevant sections of the IEEE Std 802.3-2002.

IEEE Std 802.3x A standard that defines a system of flow control for ports that operate in full duplex. 802.3x is now incorporated into the relevant sections of the IEEE Std 802.3-2002.

IEEE Std 802.1w-2001 A standard that defines Rapid Spanning Tree Protocol (RSTP) behavior.

IEEE Std 802.1X-2001 A standard that defines port-based network access control behavior.

IETF Internet Engineering Task Force. An organization responsible for providing engineering solutions for TCP/IP networks. In the network management area, this group is responsible for the development of the SNMP protocol.

IFM Intelligent Flow Management. A flow control mechanism that prevents packet loss during periods of congestion on the network.

Internet Group Management Protocol Internet Group Management Protocol (IGMP) is a protocol that runs between hosts and their immediate neighboring multicast routers. The protocol allows a host to inform its local router that it wishes to receive transmissions addressed to a specific multicast group. Based on group membership information learned from the IGMP, a router is able to determine which if any multicast traffic needs to be forwarded to each of its subnetworks.

IGMP snooping A mechanism performed by intermediate systems that optimizes the flow of multicast traffic; these intermediate systems (such as Layer 2 switches) listen for IGMP messages and build mapping tables and associated forwarding filters, in addition to reducing the IGMP protocol traffic.

IP Internet Protocol. IP is a layer 3 network protocol that is the standard for sending data through a network. IP is part of the TCP/IP set of protocols that describe the routing of packets to addressed devices.

IPX Internetwork Packet Exchange. IPX is a layer 3 and 4 network protocol designed for networks that use Novell® Netware®.

IP address Internet Protocol address. A unique identifier for a device attached to a network using TCP/IP. The address is written as four octets separated with periods (full-stops), and is made up of a network section, an optional subnet section and a host section.

Jitter An expression often used to describe the end-to-end delay variations during the course of a transmission. See also *latency*.

LACP Link Aggregation Configuration Protocol (LACP) is the IEEE 802.3ad standard. LACP provides automatic, point-to-point redundancy between two devices (switch-to-switch or switch-to-server) that have full duplex connections operating at the same speed.

LAN Local Area Network. A network of endstations (such as PCs, printers, servers) and network devices (hubs and switches) that cover a relatively small geographic area (usually not larger than a floor or building). LANs are characterized by high transmission speeds over short distances (up to 1000 m).

LLC Logical Link Control. A sublayer of the IEEE data link layer that is located above the MAC sublayer. The LLC sublayer is responsible for MAC sublayer addressing, flow control, error control, and framing.

latency The delay between the time a device receives a packet and the time the packet is forwarded out of the destination port.

line speed See *baud*.

loop An event that occurs when two network devices are connected by more than one path, thereby causing packets to repeatedly cycle around the network and not reach their destination.

MAC Media Access Control. A protocol specified by the IEEE for determining which devices have access to a network at any one time.

MAC address Media Access Control address; also called hardware or physical address. A layer 2 address associated with a particular network device. Most devices that connect to a LAN have a MAC address assigned to them as they are used to identify other devices in a network. MAC addresses are 6 bytes long.

main port The port in a resilient link that carries data traffic in normal operating conditions.

MDI Medium Dependent Interface. An Ethernet port connection where the transmitter of one device is connected to the receiver of another device.

MDI-X Medium Dependent Interface Cross-over. An Ethernet port connection where the internal transmit and receive lines are crossed.

MIB Management Information Base. A collection of information about the management characteristics and parameters of a networking device. MIBs are used by the Simple Network Management Protocol (SNMP) to gather information about the devices on a network. The Switch contains its own internal MIB.

multicast A packet sent to a specific group of endstations on a network.

multicast filtering A system that allows a network device to only forward multicast traffic to an endstation if it has registered that it would like to receive that traffic.

Network-Layer Address The network-layer address refers to a logical address that applies to a specific protocol. A network-layer address exists at Layer 3 of the OSI reference model.

NIC Network Interface Card. A circuit board installed in an endstation that allows it to be connected to a network.

OSI Open Systems Interconnection. A reference to the protocols used when interconnecting computers. These protocols relate specifically to networking and are comprised of seven layers as represented in the OSI Reference Model: physical, data link, network, transport, session, presentation, and application layer.

POST Power On Self Test. An internal test that a Switch carries out when it is powered-up.

protocol A set of rules for communication between devices on a network. The rules dictate format, timing, sequencing and error control.

Quality of Service (QoS) QoS allows the user to prioritize traffic through the network according to certain packet attributes. This ensures that high priority data is never delayed through the network by lower priority data.

repeater A simple device that regenerates LAN traffic so that the transmission distance of that signal can be extended. Repeaters are used to connect two LANs of the same network type.

resilient link A pair of ports that can be configured so that one takes over data transmission should the other fail. See also *main port* and *standby port*.

RMON IETF Remote Monitoring MIB. A MIB that allows you to remotely monitor LANs by addressing up to nine different groups of information.

router A device that provides WAN links between geographically separate networks.

RPS Redundant Power System. A device that provides a backup source of power when connected to a Switch.

Rapid Spanning Tree Protocol (RSTP) RSTP implements the Spanning Tree Algorithm and Protocol, as defined in the IEEE 802.1w standard. It provides a faster determination of network paths than the legacy STP feature.

SAP Service Access Point. A well-defined location that identifies the user of services of a protocol entity.

segment A section of a LAN that is connected to the rest of the network using a switch or bridge.

server A computer in a network that is shared by multiple endstations. Servers provide endstations with access to shared network services such as computer files and printer queues.

SLA Service Level Agreement. A "contract" between service provider and customer detailing network availability, costs, consequences of breaking the contract, and so on. The SLA may define the traffic conditioning rules or act as a framework to define the network policies. You can only configure part of the SLA using the service level or traffic aggregate table to set up your Switch to honour your network-wide 802.1p and DSCPs. You can not specify proper SLAs, which are commonly enforced by routers.

SLIP Serial Line Internet Protocol. A protocol that allows IP to run over a serial line (console port) connection.

SNMP Simple Network Management Protocol. The current IETF standard protocol for managing devices on an TCP/IP network.

Spanning Tree Protocol (STP) A bridge-based system for providing fault tolerance on networks. STP works by allowing you to implement parallel paths for network traffic, and ensure that redundant paths are disabled when the main paths are operational and enabled if the main paths fail. See also Rapid Spanning Tree Protocol (RSTP).

stack A group of network devices that are integrated to form a single logical device.

standby port The port in a resilient link that takes over data transmission if the main port in the link fails.

STP See *Spanning Tree Protocol (STP)*.

switch A device that interconnects several LANs to form a single logical LAN that comprises of several LAN segments. Switches are similar to bridges, in that they connect LANs of a different type; however they connect more LANs than a bridge and are generally more sophisticated.

Switch Database A database that is stored by a switch to determine if a packet should be forwarded, and which port should forward the packet if it is to be forwarded. Also known as Forwarding Database.

TCP/IP Transmission Control Protocol/Internet Protocol. This is the name for two of the most well-known protocols developed for the interconnection of networks. Originally a UNIX standard, TCP/IP is now supported on almost all platforms, and is the protocol of the Internet.

TCP relates to the content of the data travelling through a network — ensuring that the information sent arrives in one piece when it reaches its destination. IP relates to the address of the endstation to which data is being sent, as well as the address of the destination network.

Telnet A TCP/IP application protocol that provides a virtual terminal service, letting a user log into another computer system and access a device as if the user were connected directly to the device.

TFTP Trivial File Transfer Protocol. Allows you to transfer files (such as software upgrades) from a remote device using the local management capabilities of the Switch.

traffic prioritization A system which allows data that has been assigned a high priority to be forwarded through a switch without being obstructed by other data.

unicast A packet sent to a single endstation on a network.

VLAN Virtual LAN. A group of location- and topology-independent devices that communicate as if they are on the same physical LAN.

VLAN tagging A system that allows traffic for multiple VLANs to be carried on a single link.

VLSM Variable Length Subnet Mask (VLSM) is when the mask for a network address is different to that defined by the class of the network address.

WAN Wide Area Network. A communications network that covers a wide area. A WAN can cover a large geographic area, and may contain several LANs within it.

INDEX

Q

R

S